REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
08/18/2011	Final Technical Report	02/01/2008 - 05/31/2011

4. TITLE AND SUBTITLE
Architecture and Robust Networks

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-08-1-0043

5c. PROGRAM ELEMENT NUMBER
FA9550

6. AUTHOR(S)
John C. Doyle

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
California Institute of Technology
1200 E. California Blvd
Pasadena CA 91125

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
875 N. Randolph St
Rm 3112
Arlington, VA 22203-1954

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-OSR-VA-TR-2012 0710
FA9550-08-1-0043

12. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited A

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The aim of this research is the development of a unified theory of complex networks involving four essential aspects: the hard limits on what is achievable (constraints, misnamed —laws||), the organizing principles that succeed or fail in achieving them (architectures and protocols), the resulting high variability data and —robust yet fragile|| behavior observed in real systems (behavior, data), and the processes by which systems evolve (variation, selection, design). Hard limits on measurement, prediction, communication, computation,

15. SUBJECT TERMS
constraints, architectures, protocols, behavior,data

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	SAR	26	John C. Doyle
					19b. TELEPONE NUMBER *(Include area code)* 626 395 4808

Final Report FA9550-08-1-0043

Architecture and Robust Networks

PI: John Doyle, Caltech

PM: Dr. Robert Bonneau AFOSR/RSL (703) 696-9545
DSN 426-9545 FAX (703) 696-7360
Email: **robert.bonneau@afosr.af.mil**

Overview and background

The aim of this research is the development of a unified theory of complex networks involving four essential aspects: the hard limits on what is achievable (constraints, misnamed "laws"), the organizing principles that succeed or fail in achieving them (architectures and protocols), the resulting high variability data and "robust yet fragile" behavior observed in real systems (behavior, data), and the processes by which systems evolve (variation, selection, design).

Hard limits on measurement, prediction, communication, computation, decision, and control, as well as the underlying physical energy and material conversion mechanism necessary to implement these abstract process are at the heart of modern mathematical theories of systems in engineering and science (often associated with names such as Shannon, Poincare, Turing, Gödel, Bode, Wiener, Heisenberg, Carnot,…). They form the foundation for rich and deep subjects that are nevertheless now introduced at the undergraduate level. Unfortunately, these subjects remain largely fragmented and incompatible, even as the tradeoffs *between* these limits are of growing importance in building integrated and sustainable systems. An essential research direction then is an integrated theory based on optimization that deals systematically with uncertainty, robustness, and risk in complex systems. For a relatively nontechnical discussion of these issues, see [1] and references therein.

Insights into universal laws, architecture, and organizational principles can be drawn from three converging and increasingly related research themes. First, the organizational principles of organisms and their evolution are becoming increasingly apparent as biologists articulate richly detailed explanations of biological complexity, robustness, and evolvability that point to universal principles and architectures. Generally, organisms and their lineages are robust and evolvable in the face of even large changes in environment and system components, yet can simultaneously be extremely fragile to other small perturbations. Such universally "robust yet fragile" (RYF) complexity is found wherever we look. The amazing evolution of microbes into humans (robustness of lineages on long timescales) is punctuated by mass extinctions (extreme fragility). Diabetes, obesity, cancer, and autoimmune diseases are side-effects of physiological control and compensatory mechanisms so robust as to normally go unnoticed. This RYF feature of complex systems must be handled by any methodology that hopes to be scalable and evolvable, with systematic and formal verification approaches.

Second, while the components differ and the system processes are far less integrated, advanced technology's complexity is now approaching biology's and there are striking convergences at the level of organization, architecture, and the role of layering, protocols, and feedback control in structuring complex multiscale modularity. This complexity facilitates robustness and accelerates evolution, but enables catastrophes on a scale unimaginable without, and this "robust yet fragile" (RYF) nature of complex networks is one of their most essential features. Network-centric technology can undoubtedly provide unprecedented levels of performance, efficiency, and robustness. The ultimate challenge will not be to make this apparent in demonstrations and typical scenarios, but to avoid the rare but catastrophic real-world failures that seem to inevitably accompany new levels of complexity.

Finally, a new mathematical framework suggests that this apparent network-level evolutionary convergence within/between biology/technology is not accidental, but follows necessarily from their universal system requirements to be fast, efficient, adaptive, evolvable, and robust to perturbations in their environment and component parts ([1],[17]-[22]). The universal hard limits on systems and their components have until recently been studied separately in fragmented domains of physics, chemistry, biology, communications, computation, and control, but a unified theory is emerging ([23]-[26]). Determining what is essential about this convergence both

within biology and with technology, and what is merely historical accident requires a deeper understanding of architecture — the most universal, high-level, persistent elements of organization — and protocols. Protocols define how diverse modules interact, and architecture defines how sets of protocols are organized.

This theory builds on and integrates decades of research in pure and applied mathematics with engineering, including robust control theory, dynamical systems, information theory, numerical analysis, operator theory, real algebraic geometry, computational complexity theory, duality and optimization, and semi-definite programming, motivating new interactions between these diverse areas. The results have diverse applications, including robustness analysis of various complex control systems in biology and technology, the performance of Internet protocols and their extensions to wireless and ad hoc networks, router topologies and web layout, to wildfire ecologies, to biological signal transduction, stress response, metabolic control, and disease dynamics. The work is creating new mathematics, algorithms, and widely used software infrastructure, is appearing in the highest-impact journals in diverse fields, and concretely demonstrating that this research can help both engineers and experimental biologists. Some of the most exciting results are just recently published or are under review and not yet published, so below we will sketch some of the underlying mathematical ideas and illustrate the key results with case studies from a variety of fields.

The remainder of this report will focus on the following recent progress.

- **Generalizations of "layering as optimization" to a theory of architecture.** Our research has led to new theories of the Internet and related networking technologies (e.g [2]), and to new protocols that have been tested and deployed. We are expanding this framework to more explicitly treat dynamics ([3], [4]), and in the wireless domain both to the circuit and physical level ([5]-[10]), and to cleaner integration of routing, scheduling, power control, and network coding ([11]-[16]). Our goal is developing a common analytical framework and language that handles and integrates computation, communication, and control in complex network or networked systems across all protocol layers from physical layer to application layer and to dynamics over the network. A crucial next step is further integration with the operating systems dimensions of networking, including naming and addressing [66].

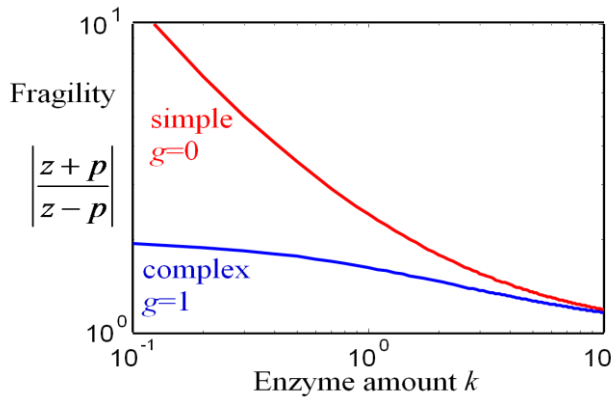


Figure 1 Hard tradeoff between net fragility (RHS of eq (5)) and metabolic overhead due to enzyme complexity and amount. Enzyme amount affects the intermediate reaction rate k (x-axis), plotted against fragility (y-axis) for g=0 (red) and g=1 (blue). Either large k or large g is required to minimize fragility, but large k requires high metabolic overhead and large g requires high enzyme complexity. Even small g>0 enhances the tradeoffs, particularly at low k.

- **New hard limits trading off robustness and efficiency**. Figure 1 illustrates a hard tradeoff between fragility of a metabolic system in terms of disturbance rejection, and the metabolic overhead of the network in terms of enzyme complexity and amount (from [24]). This appears to be the first result of its type that establishes tradeoffs between robustness and efficiency that are both theoretically rigorous and relevant to practical systems. It also gives new insights into one of the longstanding mysteries in cell biology, the purpose, if any, of glycolytic oscillations. We believe this new result is just the first in a rich set of biologically and technologically relevant hard limits. The metabolic overhead dimension is based on standard biochemistry, but is largely phenomenological and not derived directly from first principles, unlike the robustness dimension. This and other efficiency tradeoffs has motivated the next issue:
- **A fully nonequilibrium theory of statistical mechanics and thermodynamics.** We have proved abstract tradeoffs integrating communication and control theory [26], and more recently a control theoretic treatment of noise in nonequilibrium statistical mechanics [25]. This provides a theory of dissipative and active devices, the origin and nature of noise, and tradeoffs in measurement versus back action, that both integrates and generalizes standard treatments. Not yet published, but sketched below, is a "Heisenberg-like uncertainty principle" for conjugate variables that is purely classical and derivable from first principles. This research direction aims to provide a physical basis for the noise sources in control and communications, and deepen our theories on tradeoffs between robustness and efficient use of resources. We expect it will lead to a fundamental rethinking

of noise, fluctuation-dissipation, back action, and control in both classical statistical mechanics and ultimately in quantum mechanics.

- **Drag in turbulent shear flows and blunting of the turbulent profile** is a ubiquitous source of inefficiency on even highly streamlined. This has also been a long term mystery at the heart of complex systems, and one that we have largely resolved. We will sketch the essential features of our new theory.
- **Physiological variability** has been a persistent mystery at the heart of medicine from cardiology to intensive care. While not a main focus of the proposed research here, our progress in this area illustrates the power and generality of our approach, and the appendix will sketch elements of our results. We are pursuing funding elsewhere to continue this research.

Summary of Accomplishments and Research Results

From layering as optimization to a theory of architecture

The seminal work of Kelly and Low have sparked remarkable progress in mathematical modeling and analysis of the Internet congestion control, as well as later extension as a general utility maximization framework for network protocol stack design [2]. However, most theory focuses on convergence to a static optimal operating point. The duality model of TCP only says the convergence to the equilibrium rates of TCP flows, but says nothing about the transient trajectory. The dynamic nature of information over the networks and the evolution of the network itself necessitate analysis of transients and development of time-critical decision rules.

We are extending the static duality model to include transient dynamics using optimal control theory, where part of the system dynamics become a constraint on the state trajectory over time. We show that the controllers proposed by primal, dual and primal/dual algorithms all maximize some meaningful dynamical behaviors [3]. More precisely, there exist natural cost functionals whose minimization (maximization) leads to these celebrated controllers. This result opens the possibility of tackling network problems directly as optimal control problems, which not only take the dynamics into account, but which also allow to impose physical constraints. Other applications of dealing with cost functionals directly are in deducing the stability of the control system for free, gaining insight into how to perform joint routing and congestion control, etc.

The most exciting opportunity for use of the methods in [3] and others described below, however, is in more "clean slate" architectures, where control and dynamical systems theory could play an integral role at the outset, rather than patch a leaky architecture when problems (e.g. congestion collapse) arise. Thus we are rethinking the engineering network architectures (e.g. the TCP/IP protocol stack) that were the primary motivation for the development of the theory over the last decade. We have already made preliminary progress in connecting our theoretical framework with various "clean slate" efforts, such as Recursive InterNet Architecture (RINA, http://csr.bu.edu/rina/index.html), and the Publish-Subscribe Internet Research Paradigm (PSIRP, http://www.psirp.org/), and have begun new research on fundamentally redesigning network architectures.

IP has a variety of well-known problems with security, mobility and multihoming, quality of service, router table size, streaming applications, and multicasting. While the technical details are complex, many symptoms are due to the simple fact that IP addresses name *physical interfaces* (ironically, a frozen accident of the Internet's evolution). Hosts, routers, and servers have no names in IP, making multihoming and mobility difficult, while interfaces have both MAC and IP addresses. Further aggravating this is that there is no relation between IP addresses even when they are attached to the same machine (a source of persistent confusion in using traceroute to study router topology [22]). Exposing these physical addresses both globally and to higher layers is an obvious layer violation and a security, performance, and scalability nightmare. NAT (Network Address Translation) actually improves both security and scalability by creating private address spaces, but would be unnecessary in a properly layered architecture in which addresses, including physical nodes and interfaces, are local in scope and layer and properly virtualized outside.

These observations are well-known and obvious, even trivial, but they are the mere tip of an IP iceberg that netcentric technologies are obliviously plowing into. (To make matters worse, "network science" with its focus on the most elementary applications of graph theory and statistical physics is even more oblivious.) What is most interesting here is that the cell architecture solves a much harder problem than IP but has none of its weaknesses. IP networks are almost completely free of autocatalytic feedback, simply import all of their components and energy, and until recently were largely unconcerned by either waste or consumption. The cell is highly constrained and, by

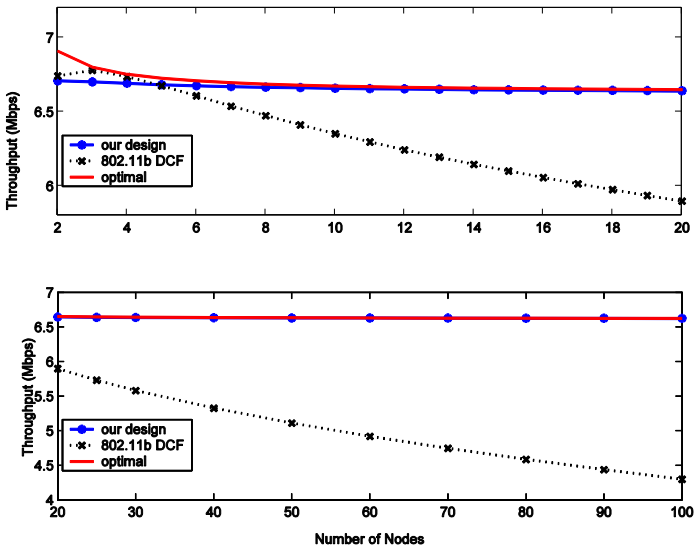
necessity, massively autocatalytic. Yet it maintains very strict layering, with specific signaling and regulatory proteins and RNAs devoted to translating the needs of the upper layering into the addresses required to access the information encoded in the genome, and a few global carriers of energy, redox, and small conserved moieties. This architecture has various scalability issues however, such as global diffusion and molecular recognition to do address resolution, and eukaryotes have more complex architectures that remain less completely understood, though the role of noncoding RNAs is clearly central and final receiving proper attention. At the other extreme, the human brain has a layered architecture that is very different from either the cell or Internet. We have already worked out qualitatively most of the many details contrasting the cell and IP-based architectures, and the next step is to formalize these observations, and connect them with our ongoing work in other projects on new Internet architectures and begin relating them to eukaryote cell and brain architectures.

Large-scale integrated circuit and antenna design

With ever-increasing demand for better performance of integrated circuits, rigorous theory-based optimal design has become more and more important, and our layering framework can be applied to this physical-layer problem as well. We have applied various ideas in the areas of passivity and convex optimization to guide the large scale linear circuit design, and showed that the problem can be cast as a decentralized control design problem for a given system [6]-[9]. As an important application of this work, the design of a secure, power-efficient, beam-steerable and on-chip transmission system for wireless networks is investigated. In particular, a passively controllable smart (PCS) antenna system is introduced, which can be programmed to generate different radiation patterns at the far field by adjusting its variable passive controller at every signal transmission. The PCS antenna is able to transmit data to a desired direction in such a way that no signal is sent in many undesired directions. Unlike the existing smart antennas whose programming leads to an NP-hard problem or are made of many active elements, the PCS antenna proposed in the present work has a low-complex programming capability and consists of only one active element. These two properties differentiate a PCS antenna from the existing smart antennas, and make it possible to implement a PCS antenna on a cheap, small-sized, low-power silicon chip.

Game-theoretic approaches to network design

Game theory has recently been applied to the analysis and design of communication networks, mainly to address self-interest and incentive issues. Most work assumes network agents (network components or users) are selfish and try to thwart selfish behaviors or induce cooperation using external mechanisms such as pricing. This economic perspective, we argue, is only one approach and by itself limits the application of a game-theoretic approach to network design. We can also take a complementary, engineering perspective to game-theoretic methods, motivated by the fact that in many network problems it is more reasonable to assume cooperative behavior. We envision a scenario where network agents are willing to cooperate but only have limited information about network states due to various practical constraints in real networks. In such a situation, the best an agent can do is to optimize some local or private objective and adjust its action based on limited information about the network state. We use a non-cooperative game to model such a situation, and let network agents behave 'selfishly' according to the game that is designed to guide individual agents to seek an equilibrium achieving some performance objective. So, from the engineering perspective, the main focus is not on incentive issues but on the implementation in practical networks. This engineering perspective complements the economic perspective and enables more flexible and practical application of game-theoretic approach to engineering design.



We have taken the engineering perspective and applied a game-theoretic approach to contention control, called random access game, which provides a unique perspective to understand existing contention based medium access control protocols and a

general framework to guide the design of new ones to improve the system performance [11]-[14]. The medium access protocol can be interpreted as and designed according to a distributed strategy update algorithm achieving the equilibrium of random access game. The random access game is a rather general construction, as it can be reverse-engineered from existing MAC protocols, forward-engineered from desired operating points, or designed based on heuristics [11][12]. Medium access methods derived from concrete random access game models achieve superior performance over the standard IEEE 802.11 DCF and can provide flexible service differentiations [12]. In the figure on the left we show the throughput comparison between our game-based medium access method, the standard IEEE 802.11 DCF, and the optimal achievable throughput. We see that our game-based design can even achieve theoretically-optimal throughput which is robust to large numbers of users.

A game-theoretic approach to distributed cooperative control has also recently received significant attention. This approach is to model the interactions of a multi-agent system as a non-cooperative game amongst the agents. The form of a distributed architecture provides advantages such as robustness to failures and environmental disturbances, reducing communication requirements, improving scalability, etc. Two main challenges of modeling a multi-agent system as a non-cooperative game are (i) designed local agent objective functions, which may be in conflict with one another and (ii) designing distributed learning dynamics so that the resulting global behavior is desirable with respect to the global objective.

However, non-cooperative control has limitations with respect to engineering multi-agent systems. One problem that arises in many multi-agent settings is coupled constraints on the agents' actions. Examples of problems that possess such constraints are consensus, formation control, or power control. We illustrate that the framework of non-cooperative games is not suitable for imposing coupled constraints. With these limitations in mind, we introduce state-based games, one particular form of stochastic games, which generalize non-cooperative games to a Markov based setting. In these state-based games, we propose an approach for dealing with coupled constraints by introducing additional information (i.e. a state) into the game-theoretic interactions [55][56].

Energy-aware network design

The use of energy has become a primary concern in system design, and computer and communication systems must make a fundamental tradeoff between performance and energy usage. The addition of energy to standard performance metrics such as delay, throughput and loss fundamentally changes the problem space of some of resource allocation designs. Not only are new mechanisms needed to optimize energy usage, existing algorithms and protocols must be re-examined as a formerly optimal algorithm may now perform poorly with respect to a new energy-aware metric. Energy management decisions must be decomposed and coordinated spatially as well as temporally, and yet global optimality must be achieved through local algorithms that are implementable in a distributed manner.

We have studied the interaction of speed scaling, a widely-adopted power management technique, with load balancing, in order to provide insights into such issues as: i) How does the system perform under speed scaling in terms of traditional performance metrics as well as energy-aware metrics? ii) How to design energy-aware optimal load balancing and can we decouple the design of load balancing from that of speed scaling? iii) How does the sophistication of speed scaling impact the design and performance of load balancing? We characterize the equilibrium resulting from the load balancing and speed scaling interaction, and introduce two optimal load balancing designs, in terms of traditional performance metric and cost-aware (in particular, energy-aware) performance metric respectively. Especially, we characterize the load-balancing-speed-scaling equilibrium with respect to the optimal load balancing schemes in processor sharing systems, and propose distributed load balancing algorithms to achieve the corresponding equilibrium and optimum. We show that the degree of inefficiency at the equilibrium is mostly bounded by the heterogeneity of the system, but independent of the number of the servers. Our results suggest that, as in many applications a low-order polynomial provides a good approximation to power function, we can decouple the design of load balancing from speed scaling without incurring much inefficiency in delay. In terms of power-aware performance metric, our results suggest that, as long as the heterogeneity in the system is small, we can decouple the design of load balancing from speed scaling without incurring much efficiency loss; but when the heterogeneity in the system is large, we have to do energy-aware load balancing if the energy consumption is a main concern [64].

Consensus in networked systems with limited channels

During the past few decades, there has been a particular interest in the area of distributed computations, which aims to compute some quantity over a network of processors in a decentralized fashion. The distributed averaging problem, as a particular case, is concerned with computing the average of numbers owned by the agents of a group. This problem has been investigated through the notion of consensus in several papers, motivated by

different applications such as flocking and synchronization of coupled oscillators arising in biophysics. We consider the consensus problem over a multi-agent network, in which the quantization effect appears due to the existence of digital communication channels between the agents. In this regard, a weighted connected graph is considered together with a set of scalars sitting on its vertices. The weight of each edge represents the probability of establishing a communication between its corresponding vertices through the updating procedure. We propose a stochastic gossip algorithm and show that the quantized consensus is reached under this algorithm for a wide range of updating parameters and any arbitrary quantizer including uniform and logarithmic ones[50]- [52]. The convergence time of the gossip algorithm is also studied. The role of these methods of reaching consensus and the control theoretic approach to protocol design will be the next direction of research.

Large-scale networked dynamical systems

One challenging issue in networked systems is the design of distributed control algorithms for spatially distributed dynamical systems. Nader Motee, as a postdoc at Caltech, has focused on tools which are suitable for development of distributed controllers for spatially distributed systems with information constraints induced by the underlying graph of the system. He has studied the locality features of distributed optimal control and optimization problems by blending tools from operator and duality theories [40][41]. In particular, it has been shown that if the coupling strength between subsystems is spatially decaying, then the large-scale receding horizon control problems can be efficiently localized in the spatial domain with stability and performance guarantees. Furthermore, this framework allows developing a method for integration and interpolation of controllers that mediates the interaction among local controllers [41]. The intent of the controller interpolation method is to produce a spatially distributed controller for a stabilizable linear spatially-varying system. In future work, we plan to build on this and other methods of Motee and Jadbabaie, which complement the approaches above.

Another issue in networked systems is the lack of a general methodology for inferring dynamical and functional behavior from the detailed network description. One of the central problems in the emerging field of systems biology is the analysis and functional classification of biochemical reaction networks. Such networks are increasingly being scrutinized and their individual components meticulously investigated in detail. Included in these large interconnected models are lists of parameter values (for component dimensions or characteristics) which are often merely known within some range or distribution. To understand how the efficiency of the entire system behaves depending on where this component parameter lies, one must perform exhaustive simulations within the parameter range. Since hundreds of parameters are often uncertain in this way, this task becomes extremely time intensive. We are exploring methodologies by which large interconnected biochemical reaction networks can be reduced to systems of much smaller state dimension that have similar functionality. The enabling ideas behind this methodology consist of understanding how dynamical systems that are designed for prescribed functions (such as logical or hybrid operations) can be implemented with dynamical networks constrained to have specific types of building blocks. This problem is referred to as functional model reduction to emphasize distinctions with traditional model reduction techniques. We propose a framework based on tools from differential ring theory and operator theory that is particularly tailored to the differential equations that result from biochemical kinetics [41][42]. This framework provides insights into uncovering and classification of function from the detailed description of biochemical reaction networks. It also proposes a new paradigm for model reduction based on network function.

Our third focus is to study networked autonomous robotic systems. Situational awareness in adversarial environments requires efficient spatio-temporal monitoring of dynamic and resource-constrained environments. A network of autonomous robotic sensors can efficiently achieve the desired spatial coverage. However, limitations on energy resources, and the required time for sensing, communication and computation place a number of hard constraints for mission planning. In addition, the robots may operate in environments cluttered with possibly moving obstacles, and their objectives can change over time. Thus, careful coordination of their paths is required in order to maximize the amount of information collected, while respecting all the constraints. While the state-of-the-art in sensor network design shows the advantages of deploying hundreds of small wireless sensors (e.g., infra-red, magnetometers, microphones, ultrasonic, and acoustic), many significant conceptual and foundational issues, such as the limits of usability and robustness of such sensor networks in urban and rural applications, remain to be studied. Our future research objectives are to further develop theories that show how to integrate mobility, computation, communication and sensing in resource-constrained mobile sensor networks, and design methodologies for distributing the computation in such systems [43]-[46].

Discrete abstractions of dynamical systems

A typical question in the integration of control and computation in complex systems would be, given a continuous control system with only finite precision measurements of the inputs and outputs, is there a finite state system that has identical input-output behavior? The associated finite state system is called a discrete abstraction of the original continuous system. Discrete abstractions are often used for automated verification and synthesis with respect to higher level temporal logic specifications. While there is already a reasonable body of work on the existence of discrete abstractions for dynamical systems, many of the current results do not give much information on the structure of the resulting finite state system, or the computational complexity of finding it. In [57], we take a known abstraction technique for discrete-time linear systems with partitioned output spaces and explicitly work out the structure of the finite state system. Properties of the output space partition control the size of the resulting finite state system, as well the complexity of finding it. In particular, with arbitrary output space partitions, the finite state spaces could grow extremely fast (non-elementary) with the order of the linear system. With simple output space partitions, however, the state spaces cannot be too large, and the discrete abstractions can be found in polynomial time. The next step in this research is to explore the implications of this result on the use of discrete abstractions in practical problems involving hybrid systems.

Control over neuron-inspired communication channels

The nervous system implements a networked control system in which the plants take the form of limbs, the controller is the brain, and neurons form the communication channels. Unlike standard networked control architectures, there is no periodic sampling, and the fundamental units of communication contain no numerical information. We proposed a novel communication channel, modeled after spiking neurons, in which the transmitter integrates an input signal and sends out a spike when the integral reaches a threshold value [54]. The receiver then filters the sequence of spikes to approximately reconstruct the input signal. It was shown that for appropriate choices of channel parameters, stable feedback control over these spiking channels is possible. Furthermore, good tracking performance can be achieved. The data rate of the channel increases linearly with the size of the inputs. Thus, when placed in a feedback loop, small loop gains imply a low data rate. Ongoing extensions of this work include analyzing a noisy version of the channel. In future work, the interplay between these channels and the layered nature of the brain will be a particular focus.

Hard tradeoffs on robust efficiency and glycolytic oscillations

Both engineering and evolution are constrained by tradeoffs between efficiency and robustness, but theory that formalizes this fact is limited. Using a simple two-state model of glycolysis, we explicitly derive hard tradeoffs between metabolic overhead, network fragility, and oscillations. These theoretical results are confirmed with single cell experiments (modeling and experiments funded elsewhere). Glycolytic oscillations are among the most studied dynamics in biology, yet whether the oscillations are beneficial or simply an evolutionary accident is unresolved. For our simple model, we prove a third alternative: Oscillations are the inevitable consequence of tradeoffs between metabolic overhead and robustness to disturbances, and the interplay of feedback control with the autocatalysis of network products necessary to power and catalyze intermediate reactions. Furthermore, the essential features of the hard tradeoff "law" depend minimally on the details of this system, and generalize to the robust efficiency of any autocatalytic network, no matter how complex. The paper [24] explores robust efficiency via integration of concepts from biochemistry and control theory [35][70] using the familiar example of glycolytic oscillations and shows how hard limits can shed new insight on a well-understood problem. We believe the main result is extremely striking and important, and is not yet published, so we will sketch the main features.

Glycolytic oscillation is a classic case study in dynamical systems, with a rich literature experimentally [71] and theoretically [72],[73]. Numerous models have been developed, from minimal models [75]-[76] to those with extensive mechanistic detail [74]. Glycolysis is arguably the most studied and most common control system, found in $\sim 10^{30}$ cells from bacteria to human, and presumably has been under intense evolutionary pressure for robust efficiency. Thus new insights are hopefully less likely to be confounded by gaps in literature or evolutionary accidents, compared with more obscure biological circuitry. Nevertheless, the purpose of oscillations, if any, is still a mystery, but one we aim to resolve in a way compatible with existing literature.

We develop the simplest possible model of glycolysis that illustrates the tradeoffs caused by autocatalysis. Other biologically motivated minimal models exist, but analysis of robustness and efficiency tradeoffs has not received much attention. Such analysis can provide a deeper understanding of the underlying basis of glycolytic oscillations. We highlight the tradeoff between steady state error and stability and its relation to the cell's metabolic overhead and pathway complexity. We then use elementary control theory to derive a more general tradeoff

formulation involving fragility, efficiency, and complexity. These deep tradeoffs show that oscillations are an inevitable consequence of metabolic efficiency and the autocatalysis. The destabilizing effects of "positive" autocatalytic feedback can be countered by negative feedback, but we show that there can be severe theoretical limits on the resulting performance and robustness. These results clarify the highly evolved nature of the yeast control network as an optimal balance of robustness, efficiency, and complexity, and are consistent with the fluctuating transient response we observe experimentally in single cells.

Minimal Model of Glycolysis

Glycolysis is a central energy producer in the cell, consuming glucose to generate Adenosine Triphosphate (ATP) used throughout the cell. Early experimental observations show that oscillations have a 90° phase difference between two synchronized pools of metabolites downstream and upstream of Phosphofructokinase (PFK) [77]. This suggests a two-state model incorporating the two Hopf modes and PFK might capture some aspects of system dynamics and indeed, such simplified models [75][76] reproduce the qualitative oscillatory behavior seen in extracts. We propose a minimal system with three reactions modeled using the power law formalism [78], for which we can identify specific mechanisms both necessary and sufficient for oscillations. (Michaelis-Menten forms complicate algebra but do not affect analysis and results, but are more familiar to biologists and the M-M form was used in [24].)

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} \triangleq \frac{d}{dt}\begin{bmatrix}x\\ y\end{bmatrix} = \underbrace{\begin{bmatrix}1\\ -q\end{bmatrix} y^{a} y^{-h}}_{PFK} + \underbrace{\begin{bmatrix}-1\\ q+1\end{bmatrix} kxy^{-g}}_{PK} + \underbrace{\begin{bmatrix}0\\ -1\end{bmatrix}(1+\delta)}_{Consumption} \tag{1}$$

We assume total concentration of adenosine phosphates in the cell [*Atot*]=[*ATP*]+[*ADP*]+[*AMP*] remains constant and the activating effects of Adenosine Monophosphate (AMP) on PFK can be modeled as ATP inhibition. ATP also inhibits PK activity. Although this seems largely ignored in most models (notable exceptions include [79]) we will emphasize its importance and model both allosteric inhibitions via exponents *h* and *g*. Since we will focus on linearizations we ignore the saturating effects of constant $[Atot]$.

In the first reaction in (1), PFK consumes *q* molecules of *y* (ATP) with allosteric inhibition by ATP. We lump the intermediate metabolites into variable *x*. In the second reaction, Pyruvate Kinase (PK) produces *q*+1 molecules of *y* at rate *k* for a net (normalized) production of 1 unit, which is consumed by the rest of the cell. We model the feedback strengths on PFK and PK as *h* and *g*, respectively, and the cooperativity of the autocatalytic reaction is modeled by *a*. In glycolysis, 2 ATP molecules are consumed upstream and 4 produced downstream, which normalizes to *q*=1 (each *y* produces 2 downstream) with kinetic exponent *a*=1. Since oscillations are typically observed in anaerobic conditions, there is no aerobic ATP production.

To highlight essential tradeoffs with the simplest possible analysis we normalize the concentration such that the unperturbed ($\delta = 0$) steady states are $\bar{y} = 1$ and $\bar{x} = 1/k$. Basal rates of PFK and consumption are normalized to 1 with perturbation δ. We focus initially on steady state error and instability using linearization, highlighting disturbance and control (the second and third term on the RHS, respectively):

$$\begin{bmatrix}\Delta\dot{x}\\ \Delta\dot{y}\end{bmatrix} = \begin{bmatrix}-k & a+g\\ (q+1)k & -qa-(q+1)g\end{bmatrix}\begin{bmatrix}\Delta x\\ \Delta y\end{bmatrix} + \begin{bmatrix}0\\ -1\end{bmatrix}\delta + \begin{bmatrix}-1\\ q\end{bmatrix} h\Delta y \tag{2}$$

The simplest robust performance requirement (motivated by the need to maintain high energy charge) is that *y* remains nearly constant despite fluctuating demand δ. This requires that the steady state error ratio

$$\left|\frac{\Delta\bar{y}}{\bar{\delta}}\right| = \left|\frac{1}{h-a}\right| \tag{3}$$

be small, or $|h-a|$ large. $\left|\Delta\bar{y}/\bar{\delta}\right| \to 0$ iff $h \to \infty$, which tradeoffs with high complexity in the first enzyme, since large *h* requires either high cooperativity or very tight ATP-enzyme binding. The resulting complex enzymes are more costly for the cell to produce. Stability of (1.2) requires $h>a$ which is consistent with (1.3) but also upper bounds the feedback strength *h*, which constrains the minimum stable steady state error to

$$\left|\frac{\Delta\bar{y}}{\bar{\delta}}\right| = \left|\frac{1}{h-a}\right| > \frac{q}{k+(1+q)g} \tag{4}$$

Equation (4) illustrates a simple and elegant tradeoff between complexity, fragility, and metabolic overhead. Low error requires large *h*, but to allow this to be stable, *k* and/or *g* must also be large. Large *k* requires either more

efficient or higher enzyme concentration and large g requires a more complex allosterically controlled PK enzyme; both would increase the cell's metabolic load. Thus low fragility directly trades off with complexity and metabolic overhead. All the tradeoffs between steady state error, complexity, and metabolic overhead achieve their hard limits when (4) is an equality, which is where (1) enters sustained oscillations (this boundary is called a supercritical Hopf bifurcation). Thus at least in this model, oscillations have no direct purpose but are side effects of hard tradeoffs crucial to the functioning of the cell.

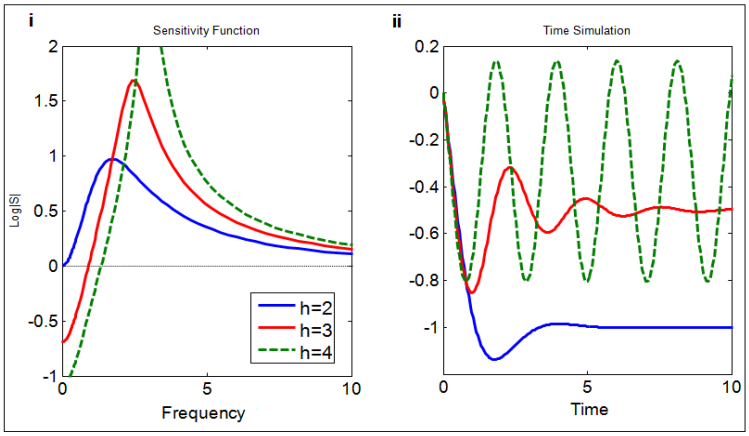


Figure 2 The log sensitivity $\log|S(j\omega)|$ without ATP feedback on PK (g=0) and the step response to change in demand δ. The integral of $\log|S(j\omega)|$ is constrained by (5) and is the same for all h. Only the shape changes with increasing h. Higher h gives lower $S(0)$ (better steady state error) with higher peak (more oscillatory transient). At q=1, the system has sustained oscillations for large h.

Hard limits on robust efficiency

Thus far we described simple tradeoffs based on basic biochemical features of a minimal model. Our elementary analysis is consistent with existing literature yet clarifies in (4) how oscillations are the inevitable consequence of robust efficiency and tradeoffs between steady state error and stability. An important next step is to add mechanistic details including more intermediates, detailed enzyme kinetics, control of redox, enzyme levels, etc, and extend the analysis to study global nonlinear stability, stochastics, and worst-case disturbances. We have extensively explored such dimensions and the results are consistent though less accessible.

A much more fundamental approach, however, is to rigorously *prove* that the tradeoffs in the simple model are truly unavoidable and independent of these neglected details; they depend only on very basic properties of autocatalytic and control feedbacks, and are neither artifacts of model simplifications nor "frozen accidents" of evolution. This will also focus our attention on the transient response to disturbances, which is just as important since even temporary ATP depletion can induce cell death [80].

We reconsider the linearized model (2) and allow $\delta=\delta(t)$ to be an arbitrary function of time. We use frequency-domain transforms for signals $\hat{y}(s) \triangleq \int_{-\infty}^{\infty} y(t)e^{-st}dt$ and transfer function $WS(s) \triangleq \hat{y}(s)/\hat{\delta}(s)$, where $s = j\omega$ are Laplace and Fourier transform variables, respectively. We consider the general case where h is replaced by a controller H with arbitrarily complex internal dynamics, constrained only to stabilize (2). Initially, H is assumed linear and time invariant and write $H=H(s)$. Given (2) and controller H, we can factor $WS(s)=W(s)S(s)$ where W is the uncontrolled (H=h=0) response from δ to y. The sensitivity function S is the primary robustness measure for feedback control [35]. $|S(j\omega)|$ measures how much a disturbance is attenuated ($|S(j\omega)|<1$) or amplified ($|S(j\omega)|>1$) at frequency ω. $S(s)\equiv 1$ when $H(s)=0$. $WS(s)=W(s)S(s)$ is the weighted sensitivity, and the response of y to any disturbance can be treated with the appropriate weight W. As shown in [23], when $q>0$, $S(s)$ has a zero where $S(z)=1$ at $z=k/q$. When $a>0$, $W(s)$ has an unstable pole ($p>0$) that is the positive real solution to $0=D(s)=s^2+(k+g+q(a+g))s-ka$ and where $W(p)=\infty$ (and $S(p)=0$)). Ideally, both WS and S should be low at all frequencies, but we can show that:

$$\frac{1}{\pi}\int_0^{\infty} \ln\left|S(j\omega)\right| \frac{z}{z^2+\omega^2} d\omega \geq \max\left\{0, \ln\left|\frac{z+p}{z-p}\right|\right\} \qquad (5)$$

with finite $z>0$ and $p>0$ as defined above (this is a variant of Bode's Integral Formula[35]). This constraint is independent on the details of $\delta(t)$, appropriate error penalties in $y(t)$ and other signals, and other sources of noise and uncertainty [70], and holds for *any* stabilizing controller H that is causal (H cannot depend on future values of $y(t)$.

H=*h* depends only on current values), no matter how complex the implementation. This "water bed" effect implies that the net disturbance attenuation ($\ln|S(j\omega)|<0$) is *at least* equaled by the net amplification ($\ln|S(j\omega)|>0$). (5) constrains $WS(s)$ for any W since W factors out. When $a>0$, $p>0$ and otherwise (5) is just bounded by 0. Hence autocatalysis always causes positive z and p and the integral in (5) is bounded by $\ln\left|\frac{z+p}{z-p}\right|$. This constraint is much deeper than the steady state tradeoffs in the previous section. Like energy and materials, robustness can be gratuitously wasted when the inequality is large, is at best conserved, and must trade off with metabolic efficiency [25]. *H*=*h* achieve (5) with equality; greater controller complexity can fine tune the shape of $\ln|S(j\omega)|$ but cannot reduce the integral.

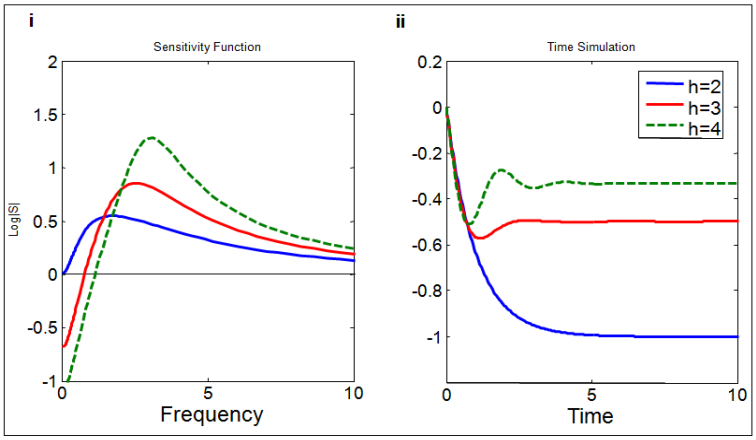


Figure 3 The log sensitivity $\log|S(j\omega)|$ and step response to change in demand δ with additional feedback loop on PK (g=1). Compared to Figure 2, both the peaks in $\log|S(j\omega)|$ and the integral are lower. Again, changing h only changes the shape but higher h is allowed without going unstable; h=4 does not drive the system into sustained oscillation as in the g=0 case in Figure 2.

The low pass filter $\frac{z}{z^2+\omega^2}$ constrains the waterbed effect to below frequency ω=z. Small z produces a more severe limitation since any disturbance attenuation must be repaid with amplification within a more limited frequency range. Since z=k/q, high k and low q are desirable. Figure 2 illustrates how autocatalysis (q=1) and (5) impact dynamics. $S(0)$ gives the steady state error while the peak in $S(j\omega)$ corresponds to how "ringy" the transient $y(t)$ dynamics are at frequency ω. At h=2, $S(0)$ is large, the peak $\|S\|_\infty$ is low, and $y(t)$ has a large steady state error, which h=3 lowers but with more transient fluctuations. At h=4 the system oscillates at the frequency where $S(j\omega)\to\infty$. The tradeoff in (4) disappears and the bound in (5) $\to$0 with no autocatalysis ($q\to$0). Zero steady state error with stability is then possible by taking $h\to\infty$.

Models and experiments revisited

We argue that PK feedback plays an important role in stabilization. Increasing g decreases p (leaving z unchanged), decreasing $\ln\left|\frac{z+p}{z-p}\right|$, uniformly improving constraint (5) and the stability bound in (4). If q=a=1, the system is stable for *all* k>0 iff $0<h-1<2g$. Thus g>0 is necessary to simultaneously maintain acceptable steady state error $S(0)=1/(h-1)$ and stability for all k>0. Replacing g=*0* (Figure 2) with g=1 (Figure 3) doesn't change $S(0)$, but the peak and integral of $\ln|S(j\omega)|$ are lower and $y(t)$ is more damped. h=4 is unstable in Figure 2 but stable in Figure 3

Much more significant is the effect g>0 has on the robustness vs. efficiency tradeoff involving k. While a and q are essentially fixed by the network's autocatalytic structure, h and g can be tuned on evolutionary time scales. Thus $0<h-1<2g$ is biologically plausible and in fact consistent with most estimates, ensuring stability for all k>0. This allows individual cells to fine tune k>0 via the myriad mechanisms that control enzyme levels. Since z=k/q, increasing k improves both sides of (5) and uniformly improves robustness, at the expense of higher enzyme levels and thus higher metabolic overhead (Figure 1). Stability for all k>0 also relates to robustness to noise in gene expression and enzyme levels, though quantifying this effect would require more detailed modeling which we intend to pursue. From an engineering perspective, this is a remarkably clever control architecture, and the presence of g>0 suggests that at least in this case evolution favors higher complexity in exchange for this kind of flexibility and robustness. However, expanding our simplistic model mostly introduces other fragilities, so the possibility of single

cell oscillation cannot be ruled out theoretically or experimentally. Explicitly modeling and understanding autocatalytic and control feedback of redox via NADH is the next most obvious source of further hard limits.

The relationship between the above analysis and experiments is subtle, but consistent. From the hard constraints (4)-(5)we can easily identify worst case conditions. Small $z=k/q$ increases overall fragility. This occurs at high autocatalytic stoichiometry q, most easily created by anaerobic condition since there is no ATP production from aerobic metabolism. But low intermediate reaction rate k has a similar effect, so our experiments also aim for conditions that might give low k, including growth in ethanol and amino acid starvation (experimental evidence shows decreased level of some glycolytic transcripts when *S. cerevisiae* is grown in ethanol [82], which could decrease k.) Cells were then shifted into anaerobic glucose metabolism. Our model does not predict sustained oscillations, but interesting transient dynamics are possible as is some cell to cell variability due to enzyme level variations [83]. Single cell NADH autofluorescence showed no sustained oscillations, but a portion of the cells exhibited fluctuating transients before settling into higher NADH level (Figure 4). The period is in good agreement with the 36s period in cell suspensions [81].

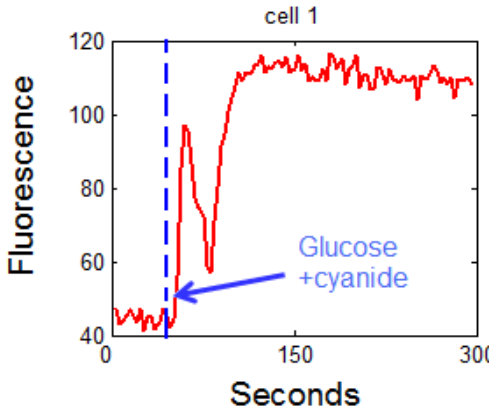


Figure 4 Single cell trace of NADH autofluorescence in previously-starved yeast cells made anaerobic using potassium cyanide (KCN). Dashed line indicates when the media is switched. A portion of the cells exhibited fluctuating transient before settling into a higher NADH level. The period of the fluctuation transient is in good agreement with the period of sustained oscillations in intact cells reported in [81].

Implications of hard tradeoffs and glycolytic oscillations

Our analysis illustrates the power of control theory to clarify biological phenomena, and biology to motivate new theoretical directions [84]. In this simple model of glycolysis, oscillation is neither directly purposeful nor an evolutionary accident but a necessary consequence of autocatalysis and hard tradeoffs between fragility, efficiency, and complexity. Nature has evolved a feedback structure that effectively manages these tradeoffs with flexibility to adapt to changes in supply and demand, at the cost of higher enzyme complexity. Consistent with engineering, complexity in biology is primarily driven by robustness, not minimal functionality [19].

Our theory is consistent throughout in highlighting hard tradeoffs, but there are important differences in details. While (4) is phenomenological and specific to the model in (2), (5) is much deeper. It holds for arbitrary causal controllers no matter how complex, and applies to other systems. However, (5) still requires substantial phenomenology since the formulas for z and p depend on assumptions about autocatalysis (q and a) and enzyme efficiencies and levels (k). This motivates further unification of control theory with thermodynamics and statistical mechanics. Recent progress is encouraging [25], and we will consider this direction next. It also motivates rethinking how biology overcomes the "causality" limit with various mechanisms that exploit predictable environmental fluctuations (e.g. circadian rhythms) or provide remote sensing (e.g. vision, hearing), both of which can greatly mitigate hard limits such as (5) [26]. In the case of circadian rhythms, oscillation is not just a side effect but has the purpose of exploiting the cyclic environment.

To make our ideas accessible, we used the simplest possible model that captures the real system's essential features yet facilitates theoretical analysis connecting network structure with functional tradeoffs. We have extended these models in various ways, described in the appendix. While our minimal model has limited quantitative predictive power, it can still provide qualitative insights about experiments. Our approach restricts the controller implementation using ATP inhibition. Allowing for arbitrary control by any intermediate will change the network's feedback topology which, as in the case of PK inhibition, seems to be more effective at lifting stability and performance constraints at the cost of pathway and enzyme complexity. A few glycolytic enzymes are inhibited by their immediate products, again suggesting that nature favors higher complexity to gain robustness.

Measurement Limitations and Statistical Mechanics

To replace the phenomenology in enzyme amount and complexity in Figure 1 with first principles theory of efficiency, we need a rigorous treatment of enzymatically catalyzed reactions that are both allosterically controlled and far from thermodynamic equilibrium. Similar challenges will arise in any theory connecting robustness with efficient use of energy and materials. Unfortunately, standard methods in thermodynamics and statistical mechanics were never developed for such purposes and are clearly inadequate and incomplete, a situation we have begun to rectify. In [25] we take a control-theoretic approach to answering some standard questions in statistical mechanics, and use the results to derive limitations of classical measurements. We will illustrate the results in [25] as applied to the simplest possible measurement problem and also extend them to some striking new limits. Consider a simple abstract one state linear stochastic differential equation

$$\begin{aligned} \dot{v} &\triangleq \frac{dv}{dt} = -\frac{1}{m}\sqrt{\frac{2kT}{R}}w \\ y &= v + \sqrt{2kTR}w \end{aligned} \tag{6}$$

Here the state v can be thought of as the velocity of a particle of mass m (or the voltage across a circuit with capacitance m and resistance R.) The output y is assumed to be from a sensor of with friction R. A main result of [25] is that even highly idealized sensors must have a white sensor noise w with unit intensity as in (6), where k is Boltzmann's constant and T is the sensor temperature. There is also a corresponding stochastic back action as in (6). While this is consistent with standard phenomenological arguments, [25] provides new, rigorous, and elegant derivations.

A central problem is the relation between systems which appear macroscopically dissipative, such as those having resistance and friction, but are microscopically lossless. We show in [25] that a linear system is dissipative if, and only if, it can be approximated by a linear lossless system over arbitrarily long time intervals. Hence lossless systems are in this sense dense in dissipative systems, a particularly elegant resolution of the origin of dissipation. A linear active system must be approximated by a nonlinear lossless system that is charged with initial energy. While the combination of nonlinearity and energy is often thought to cause unpredictable dynamics, even chaos, in this setting, nonlinearities are essential resources, like energy, in building nontrivial but organized systems [1]. These distinctions are central to our whole approach to complex networks.

As a by-product of these results, we obtain mechanisms explaining the Onsager relations from time-reversible lossless approximations, and the fluctuation-dissipation theorem from uncertainty in the initial state of the lossless system. The results are applied to measurement devices and are used to quantify limits on the so-called observer effect, also called back action, which is the impact the measurement device has on the observed system, as in (6). In particular, it is shown that deterministic back action can be compensated by using active elements, whereas stochastic back action is unavoidable and depends on the temperature of the measurement device. Hence the form of (6) above. Also, the measurement in (6) requires an active component to zero the deterministic back action of the sensor, which is assumed here to have an unlimited energy source ([25] also shows how limited energy to the sensor adds further to the hard limits described here).

Now consider the problem finding an estimate $\hat{v}(t)$ of the state $v(t)$ to minimize $E\left(\hat{v}(t)-v(t)\right)^2$, assuming no knowledge of the state at t=0, i.e. that $v(0)=v_0$ is unknown and $E\left(\hat{v}(0)-v(0)\right)^2=\infty$. This is a standard Kalman filtering problem that because of its special structure can be solved analytically, and the resulting optimal values have particularly simple formulas (with admittance Y=1/R)

$$\left.\begin{aligned} E(v-v_0)^2 &= \frac{2kTt}{Rm^2} \\ E\left(v-\hat{v}\right)^2 &= \frac{4kT}{\left(1-e^{-2Yt}\right)} \geq \frac{2kTR}{t} \end{aligned}\right\} \Rightarrow \sqrt{E\left(v-\hat{v}\right)^2 E\left(v-v_0\right)^2} = \frac{2kT}{m}\left(\frac{2t}{R\left(1-e^{-2Yt}\right)}\right)^{\frac{1}{2}} \geq \frac{2kT}{m} \tag{7}$$

This extremely simple result shows that there is a hard tradeoff between measurement error and back action that depends on T and m. Not surprisingly, all quantities are worsened by high sensor T, and the back action and product worsened by small mass m. The product is optimized at small time t and t/R trades off error with back action. The intuition behind (7) is also clear, in that measurement allows for rapid reduction in estimation error with diminishing returns over time, while the resulting back action causes the velocity to undergo a random walk that worsens over

time. A natural strategy would be to make a near instantaneous measurement on a small interval $[0,\tau]$, using R to balance estimation error with back action, then disconnect the sensor and propagate the optimal estimate, which without measurement is constant, since $\frac{d\hat{v}}{dt}=0\Rightarrow\hat{v}(t)=\hat{v}(\tau)\,\forall t\geq\tau$ with $E\left(v(t)-\hat{v}(t)\right)^2\approx\frac{2kRT}{\tau}\,\forall t\geq\tau$. Note that between measurements, there is no difference between "plant" and filter dynamics, and that during a measurement the error $E\left(v-\hat{v}\right)^2\propto\frac{1}{t}$ undergoes a "collapse." Thus even in this purely classical setting, careful accounting for noise and back action using the results in [25] and standard methods of control theory prove some surprising consequences of a type usually associated exclusively with quantum mechanics. The tradeoffs involving the sensor energy supply are equally interesting but the derivations more involved [25].

The well-known Heisenberg uncertainty principle in quantum mechanics does not involve plant versus filter states (physics makes no such distinction), but between position and momentum. To explore analogous classical properties, assume the particle has position x with dynamics $\dot{x}=v$ and hits a highly idealized detector at x=0 and t=0, with velocity and sensor dynamics as in (6). Assume this position is known perfectly at t=0, but velocity is again unknown, and a measurement and optimal estimator is performed on a small interval $[0,\tau]$. This too can be solved analytically, with velocity and error estimate as in (7) but with position errors

$$E\left(x-\hat{x}\right)^2=\frac{4kT\left(1-e^{-Yt}\right)^2}{Y^2\left(1-e^{-2Yt}\right)}\geq\frac{2kT}{Y}t\Rightarrow\sqrt{E\left(v-\hat{v}\right)^2E\left(x-\hat{x}\right)^2}=\frac{4kT\left(1-e^{-Yt}\right)}{Y\left(1-e^{-2Yt}\right)}\geq\frac{2kT}{Y}=2kTR \tag{8}$$

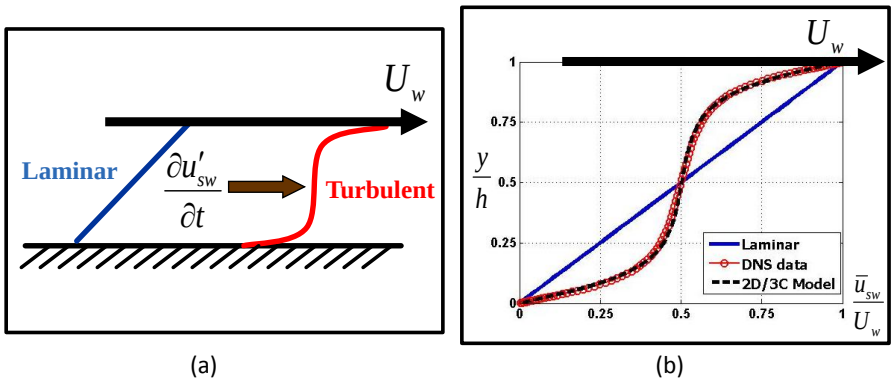
The "uncertainty principles" in (7) and (8) show a tradeoff between error estimates for velocity (or momentum) and either velocity back action (7) or position estimate (8). As noted above, neither classical nor quantum physics makes a distinction between plant and filter states, whose dynamics are identical between measurements. If measurements are assumed to be nearly instantaneous, their main effect is to "collapse" the estimates of the measured variables, subject to the limits in (7) and (8).

We believe the approach in [25] is just the beginning of a fully nonequilibruim theory of noise, fluctuation-dissipation, back action, measurement, and control, and ultimately tradeoffs between robustness and efficiency. We expect it will force a fundamental rethinking of these issues both in a classical setting but ultimately in quantum theory as well. What the hard limits in (5) above have in common with (7) and (8) is that feedback and dynamics are powerful tools (as are nonlinearities) but their implementation involves hard limits on achievable performance and robustness. Since limits in both communications and control theory depend on abstract notions of sensor and channel noise, a natural next step is to use the methods in [25] to trace the implementation tradeoffs in both domains to physical mechanisms involving dissipative and active devices and their energy requirements.

Perhaps more profoundly, existing theories of both quantum mechanics and classical statistical mechanics both lack a complete measurement theory of the type in [25], the essence of which involves simply assuming that the measurement devices must be implemented with the same physics as the "plant". A natural next step then is to generalize [25] to the problem of measurement of quantum phenomena but implemented physically with devices that must also obey quantum dynamics. Mathematically, the aim is to replace the measurement axiom with a theorem. Just as [25] is consistent with existing near-equilibrium theory, a rigorous theory of quantum measurement would not affect the predictions for standard experiments, but could have profound implications for engineered systems exploiting active devices and far-from-equilibrium behavior. Thus a principle aim of our research is to connect these mathematical issues with more physically-based critiques of the conventional theory, such as [67], which has largely been ignored or dismissed both in engineering and physics.

Turbulent profiles and drag

Turbulence is another of the persistent and unsolved problems in physics, but also essential to efficiency in engineered systems. There are many unresolved issues, however in wall bounded shear flow the shapes of the laminar and turbulent velocity profiles are well known. Those for plane Couette flow are depicted in Figure 5 (a). However, the underlying mechanisms involved in creating the S shaped (blunted) turbulent profile remained unresolved. Understanding this phenomenon may be the key to developing control strategies to prevent transition to turbulence in applications where the associated increases in drag cause undesirable effects.



(a) (b)

Figure 5 (a) 2D depiction of the flow configuration and the laminar and turbulent velocity profiles for plane Couette flow with a moving upper plate. (b) The results from the 2D/3C simulation capture the profile blunting.

Linear models can generate flows dominated by the streamwise elongated structures that are prevalent and play an important role in fully developed turbulence. However, a nonlinear model is required to capture the momentum transfer that produces a turbulent velocity profile. Numerical and experimental observations of these streamwise and quasi-streamwise elongated structures motivate the study of a streamwise constant projection of the Navier Stokes equations. The resulting two-dimensional, three-velocity-component (2D/3C) nonlinear model captures important nonlinear features of turbulence, while maintaining the linear mechanisms that have been shown to be necessary to maintain turbulence.

In our work we attempt to rigorously connect the observed flow features to the creation of the turbulent mean velocity profile. We show that in a robust control framework, the so-called 2D/3C model captures the blunting of the profile along with other salient features of fully developed turbulent plane Couette flow [31]. The robust control framework employs small amplitude Gaussian noise forcing to simulate the 2D/3C model's response in the presence of disturbances, uncertainty and modeling errors. Figure 5(b) shows the mean velocity profile obtained from the simulation compared to experimentally verified direct numerical simulation (DNS) data. The model captures the change in mean velocity profile from the nominal laminar to the characteristic "S" shaped turbulent profile [29]. Comparison of the full velocity field with a spatial field of DNS demonstrates that the simulations also capture the salient features of fully developed turbulence [27], as seen in Figure 6.

The laminar flow solution is globally stable, which indicates that the flow state should always return to the streamlined laminar flow condition. The fact that the 2D/3C model is able to generate "turbulent-like" behavior under small-amplitude stochastic noise indicates that transition to turbulence in this model is likely a consequence of the laminar flow solution's lack of robustness (inability to maintain the maintain this flow condition) in the presence of disturbances and uncertainty [30]. In fact, large disturbance amplification is common in both this model and

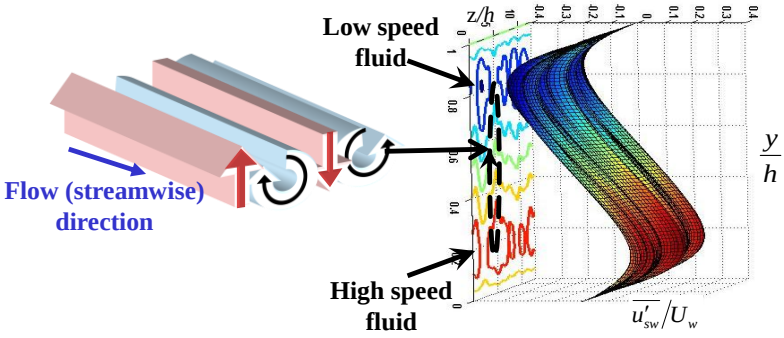


Figure 6 Cartoon of streamwise elongated structures with surface plot of 2D/3C simulation results

the linearized Navier Stokes equations.

A steady-state (time-independent or static) version of the 2D/3C model is employed to isolate the mathematical mechanisms that are involved in generating an appropriately shaped turbulent velocity profile [28]. We use cross-stream components (i.e. those representing a cross-section of the three-dimensional flow) to create a model of structures consistent with the experimentally and numerically observed flow features. This is used as input to develop a forced 2D/3C streamwise velocity equation, i.e. we study a model of the velocity component that describes the shape of the mean profile. The resulting steady-state solutions are shown to have the same qualitative features as both a spatial field of DNS data and the results of the full stochastic simulations. These results provide evidence that the nonlinear terms in the 2D/3C streamwise velocity equation are responsible for the momentum transfer associated with the change in profile from the nominal laminar to the turbulent state. Analytical study of the equations confirms that the momentum transfer that produces the correct mean profile requires a nonlinear momentum equation for the streamwise velocity component. Isolating these momentum transfer mechanisms represents an important step in the development of flow control strategies because delaying the onset of turbulence and turbulence suppression are common goals in flow control applications.

Finally we attempt to make a connection to between the linear mechanisms involved in large disturbance amplification and the nonlinearity required for the blunting [28]. We find that the linear equations allow us to appropriately model the width (spanwise extent) of the streamwise coherent structures. However, this comes at the expense of capturing the mean velocity profile. There appears to be an important tradeoff between these two mechanisms. These types of tradeoffs are very common in engineering systems and understanding them provides important information in designing systems that are both safe and able to meet advanced performance requirements. Further understanding of this tradeoff may also provide important insight into the mechanisms associated with both transition and fully turbulent flow.

Appendix (selected details and extensions)

Glycolytic oscillations

Model Extension and Topological Effects

In [34], we show the effects of the pathway size, reversibility of intermediate reactions and consumption of intermediate metabolites on performance. In addition, we establish some necessary conditions on the existence of fixed points and their stability. We show that for the general model there exist lower and upper bounds on the feedback gain that guarantee stability for arbitrary pathway size as well as arbitrary values of both the (reversible) reaction rates and the intermediate consumption rates. These bounds are tight in the sense that for gains that lie outside the ranges established by these bounds, we can construct specific unstable pathways.

For the general model we again show that increase in the intermediate reaction rates makes the pathway more stable, increases the magnitude of the RHP zero, and softens the hard limits on performance. Increased pathway size has the opposite effect. An increase in the size of the chain of enzymatically catalyzed intermediate reactions in the autocatalytic pathways causes two adverse effects on performance: tradeoffs on performance limits are exacerbated (as RHP zero becomes smaller) and the range of available stable gains is reduced, which makes the operating gains less robust and reduces the achievable performance objectives.

Consumption of intermediates in autocatalytic metabolic pathways results in less resources available to convert to the product of the pathway, which effectively reduces the net product of the pathway. This effective reduction in net output production makes the pathway harder to control because it corresponds with the RHP zero getting smaller, and thus aggravating the tradeoffs on performance limits. Additionally, the pathways must enforce a positive return in energy investment, therefore the loss of the intermediates to the other pathways must be only a fraction of the available resources, otherwise excessive consumption of these intermediates causes the pathway to crash. On the other hand, the presence of reversible reactions makes autocatalytic pathways easier to control, as they function as “release valves" by making higher stable gains available, thus providing more robustness and better achievable performance objectives.

Nonlinear Results

This paper focuses entirely on linearizations but we have extended our model to a nonlinear model of arbitrary length. In [1] we discuss an approach based on system-theoretic measures, such as the extent of region of attraction (RoA) around the nominal operating points of the system, to prove robustness under initial conditions perturbations. We demonstrate the use of the approach on a specific class of autocatalytic pathway models that capture the core structure of the glycolysis pathway.

In [32]we show that the size of the estimated (through a numerical optimization-based procedure) RoA around the nominal operating condition provides information about the robustness of the model to parameter perturbations. More specifically, numerical experiments demonstrate that systems that are robust with respect to perturbations in the parameter space have large, easily “verifiable" (in terms of proof complexity) estimates of the RoA. Additionally, for systems close to the stability boundary, small changes in the feedback strength lead to several different regimes in which “simple" polynomial Lyapunov functions (i) certify large invariant subsets of the RoA; (ii) can only certify relatively smaller sets to be in the RoA; (iii) cannot certify (to the tolerances used in the numerical computations) any invariant subset of the RoA. This optimization-based procedure becomes computationally impractical as the pathway size increases. In order to extend the RoA analysis to larger pathways, we take a compositional approach which exploited a natural decomposition of the system, induced by the underlying biological structure. The pathways are decomposed into a feedback interconnection of two input-output subsystems, a small subsystem with complicating nonlinearities and a large subsystem with simple dynamics. This decomposition simplifies the analysis by assembling RoA certificates based on the input-output properties of the subsystems. The simplest decomposition allows us to analytically construct, using storage functions and simple quadratic supply rates, block-diagonal Lyapunov functions for a large family of autocatalytic pathways. We show that if a Lyapunov function of the specified block-diagonal form exists, then it can be constructed using this decomposition. For analysis of a larger class of pathways, more general versions of the decomposition are required, allowing for the size of the subsystem with the complicating nonlinearity to increase. This strategy leads to two conflicting trends: a larger family of pathway models becomes amenable to RoA analysis at the expense of computational complexity.

Physiological Variability in Health and Disease

For decades, research has shown that variability in a physiological signal (e.g. heart rate, respiration, blood pressure) is associated with the status of the cardiopulmonary and autonomic nervous systems. A recent direction in our research is to apply rigorous system identification and control theoretic tools to the analysis and modeling of heart rate variability (HRV), including connecting data with mechanistic physiological models. We have used these tools to understand the responses of healthy, fit human subjects to exercise, both to drive tool development and establish a deeper understanding of mechanisms involved in healthy HRV. Our plan is to apply these system identification and control theoretic tools, along with insights on healthy HRV, to gain mechanistic insights into developing alert algorithms for patient monitoring and diagnosis. This is the most radical and potentially controversial dimension of our proposed work, but also has the highest potential *clinical* impact. While we are primarily seeking funding for this from other sources, the results have potentially huge implications for DOD, and also illustrate the breadth of our tools, so we will briefly review some of the basic ideas in this research direction.

The dominant theme in physiological signal variability analysis has been the use of the tools of time series statistics, chaotic dynamics, and statistical physics to assess the well-being of the cardiopulmonary system. However, this approach ignores the mechanisms causing the variability in a signal, while only examining the statistical behavior of a signal in isolation. This has the following weaknesses in terms of the properly engineered development of an effective clinical alert for control dysfunction: 1) It has generally ignored the important fact that the observed signal variability might be due to the influences of other signals. Our current results successfully model the causal response of heart rate (HR) to watts and ventilation during exercise. This evidence supports the idea that physiological mechanisms can actually be discerned if we employ the correct mathematical tools and domain knowledge. 2) Statistical analysis alone can find the correlation but not the *causality* between signals. Correlation at best tells us how we might detect the occurrence of an impending "crash" of a system. It does not report the development of abnormalities in the control system itself that might provide an adequate time window to prevent a crash of the system under control. 3) Statistical analysis does not naturally deal with homeostasis. Homeostasis is not the simple loss of signal variability but rather is the manifestation of working controls (heart rate, ventilation…etc), that act, sometimes aggressively, to minimize errors (blood pressure, pH, O2 saturation,…etc). 4) Statistical analysis requires the system to be stationary and fails to give proper interpretation in non-stationary ('dynamic') conditions, which are universal in a real-world clinical environment. 5) Statistical analysis fails to accurately identify the physiological mechanism(s) underlying physiological signal variability. A control theoretic framework can in principle solve all of the above problems but will require a substantial shift in thinking towards dynamics and homeostasis, the role of physiology, and mechanistic interpretation of signals.

In our framework, we propose to combine a black-box model (system ID) with a mechanistic model in order to link signal variability and physiological mechanism. This linkage may eventually provide clinicians with useful information for diagnosis, prognosis and treatment as follows: An initial model in healthy subjects will help us characterize the fundamental physiology. To some extent, the differences between athletes and normal healthy people can be regarded as analogous to those between healthy people and those with cardiopulmonary malfunction. While it is very important to understand the control of physiological parameters, per se, to a clinician [68] the most exciting possibility raised by this kind of deeper understanding is detection of clinical problems at *earlier stages than is now possible*. For example, while the raw simple monitored data may not clearly reflect a developing anomaly, control system analysis may do so. This introduces an important new concept in pathophysiology: Detection of control system deterioration or dysfunction that precedes failure of the actual organ system. Such detection may provide an upstream mechanism that allows for more proactive intervention and less subsequent downstream damage. Even if this downstream damage is not completely preventable, monitoring systems incorporating this knowledge could provide alerts in a more remediable phase before irreversible structural damage is done to key components. In addition, the control element analyses might be combined and placed in clinical context to formulate even more sensitive and accurate alerts.

Heart Rate Variability

Human heart rate signals exhibit a high degree of variability. Reductions in heart rate variability (HRV) are often associated with disease**,** and so an understanding of the underlying mechanisms is critical [86] [88]. It has been proposed that heart rate variability arises from stochastic, nonlinear, and possibly chaotic dynamics [88]- [92]. Results from mathematical tools such as time domain, fourier, wavelet and multifractal analysis as applied to isolated heart rate signals have been the basis for such insights [92] -[95]. These methods, however, have two shortcomings: nonstationarity in signals due to background influences are not naturally handled, and the

physiological mechanisms underlying fluctuations in heart rate are not addressed. We have adopted two dynamical modeling approaches to overcome these shortcomings and explain the source of heart rate variability. We used 'black-box' techniques [101] to directly model deterministic causal relationships between exogenous disturbances and heart rate signal, and we used 'first-principles' models to quantify plausible physiological mechanisms that correspond to these relationships. We find that slow time scale variation in heart rate during dynamic exercise can be captured by simple linear models governing heart rate response to dynamic workload. Fast time scale variation at fixed workloads can be captured by simple linear models governing heart rate response to ventilation. Moreover, these simple models allow us to characterize simple nonlinear (but not chaotic) dynamics exhibited in heart rate response to simultaneous excitation of ventilation and workload.

Figure **7** shows three cycling experiments that strikingly illustrate HRV. (Qualitatively similar results hold for other fit, healthy subjects, but with significant quantitative variations both between subjects and over time.) The HRV at low watts is dramatically higher on both slow and fast time scales. Such HRV is widely believed to be a signature of health, and its loss a symptom of disease. Thus at least superficially, the reduction in HRV with increasing watts levels in healthy athletes mirrors this well-known "signature" change in HRV from health to disease, but in both cases detailed mechanisms underlying this change remain murky. Also shown are the output h=HR of $\Delta h(t) = h(t+1) - h(t) = ah(t) + bw(t) + c$, a simple local linear model with w=watts input and constants (a, b, c) fitted to minimize the error between $h(t)$ and HR data. No simple model with similar error exists that does not include watts inputs, so the large slow fluctuations are consistent with the obvious and well-understood need of cardiovascular control to meet changing watts demands. A single global model for all watts levels would necessarily be nonlinear, since the parameter values (a, b, c)≈(-.1, .1, 7) @ 0w differ greatly from (-.02, .01, 1) @ 100w. This is further confirmed by simulating (in blue) HR with the model fitted for the middle exercise but with easy and hard exercise as inputs. (A single simple nonlinear model can indeed fit as well as the three separate linear models).

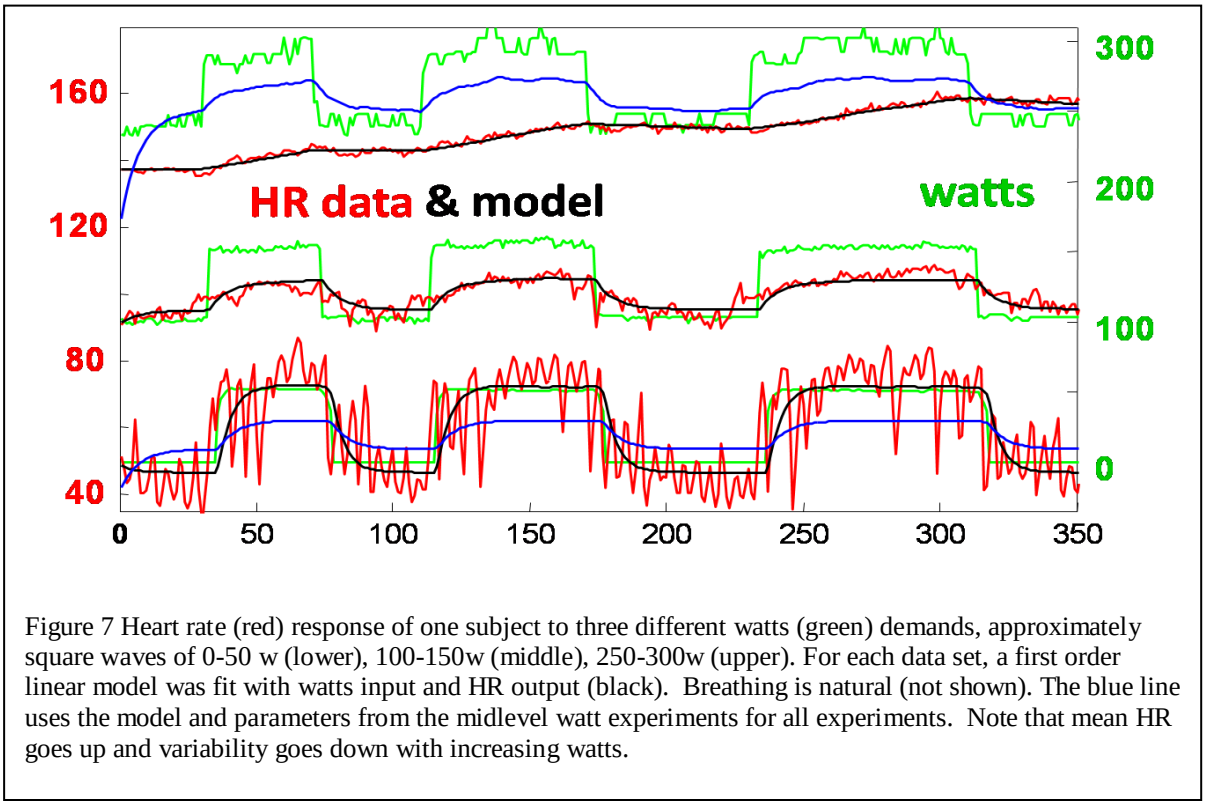


Figure 7 Heart rate (red) response of one subject to three different watts (green) demands, approximately square waves of 0-50 w (lower), 100-150w (middle), 250-300w (upper). For each data set, a first order linear model was fit with watts input and HR output (black). Breathing is natural (not shown). The blue line uses the model and parameters from the midlevel watt experiments for all experiments. Note that mean HR goes up and variability goes down with increasing watts.

To investigate mechanisms for both the linearized dynamics and their nonlinear changes with watt levels we consider a standard first-principles model of aerobic cardiovascular control [98][100][105] , with blood and oxygen circulation, peripheral vasodilatation and increased oxygen consumption induced by exercise, and the effect this and heart rate has on blood pressure and oxygen saturation. Assuming that systemic arterial O2 is controlled by ventilation allows removal of both from the model, yielding

$$\begin{aligned} c_{as}\dot{p}_{as} &= c_l \cdot H \cdot p_{vp} - (p_{as} - p_{vs})/R_s & c_{vp}p_{vp} &= V_{total} - \left(c_{as}p_{as} + c_{vs}p_{vs} + c_{ap}p_{ap}\right) \\ c_{vs}\dot{p}_{vs} &= (p_{as} - p_{vs})/R_s - c_r \cdot H \cdot p_{vs} & V_{t,O_2}[\dot{O}_2]_t &= -M_{o_2} + F_s \cdot ([O_2]_a - [O_2]_v) \\ c_{ap}\dot{p}_{ap} &= c_r \cdot H \cdot p_{vs} - (p_{ap} - p_{vp})/R_p & \dot{H} &= u(\cdot) \end{aligned} \quad (9)$$

$[O_2]_t$ is the tissue oxygen saturation. The *p* variables are pressures with subscripts *a*=arterial, *v*=venous, *s*=systemic, and *p*=pulmonary. At steady state we can solve for blood pressure and oxygen saturation as a function of heart rate and watts, (BP, ΔO2) = f (HR, W), where BP is the mean systemic arterial blood pressure and ΔO_2 is $[O_2]_a$-$[O_2]_t$. The right mesh plot in Figure 8 is the image on the BP-ΔO2 plane of the left HR-W mesh plot under the function f (HR, W). The solid curve shows idealized but typical steady state values of heart rate as a function of watts, and its effect on (BP, ΔO2).

Of course, the implementation of the autonomic nervous system's control of heart rate serves as a proximate cause for decreasing HRV with increasing watts. It is known that HRV is directly correlated with parasympathetic tone and that parasympathetic stimulation is inhibited by sympathetic tone [106]- [109]. Thus, larger watt levels imply larger sympathetic tone which implies decreased HRV. Just as with glycolytic oscillations the ultimate cause, however, remains: *why* is the nervous system implemented this way? We aim to explain both Figure 7Figure 8 in terms of familiar physiological tradeoffs, roughly analogous to the tradeoffs that were explored in the context of glycolysis but involving different variables. To start with, a hypothetical linear response (solid to dotted line) consistent with the low watts data can be explained in terms of purely metabolic tradeoffs and brain environment homeostasis. With proper hydration, nutrition, and sleep, healthy subjects can maintain moderate watts levels almost indefinitely. This requires relatively high HR to maintain high tissue O2 (low ΔO2) and maximize aerobic lipid metabolism, preserving precious carbohydrate energy sources, and this can be done with modest metabolic overhead for HR itself.

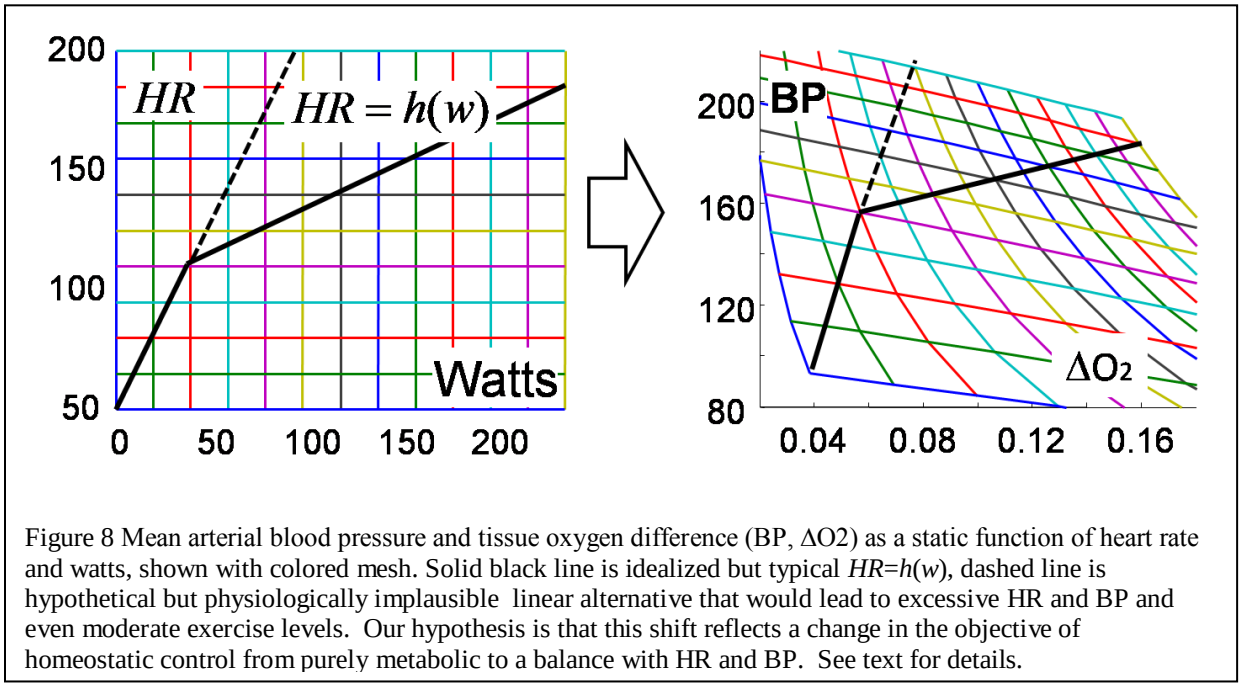


Figure 8 Mean arterial blood pressure and tissue oxygen difference (BP, ΔO2) as a static function of heart rate and watts, shown with colored mesh. Solid black line is idealized but typical *HR*=*h*(*w*), dashed line is hypothetical but physiologically implausible linear alternative that would lead to excessive HR and BP and even moderate exercise levels. Our hypothesis is that this shift reflects a change in the objective of homeostatic control from purely metabolic to a balance with HR and BP. See text for details.

The actual nonlinear response in Figure 8 (solid line) reflects additional tradeoffs. For typical maximum heart rates, the dashed line is not achievable for even modest watt increases. In addition, at high watts and HR, blood pressure would be elevated to levels that are potentially damaging, while in fit athletes there is diminishing benefit of high tissue O2 (low ΔO2) because muscle mitochondria saturate. All these factors can be quantitatively reflected in a static least squares optimal control model of $u(\cdot)$ by assuming that penalties on BP and HR relative to ΔO2 increase with watts, and simple computations do reproduce the typical steady state values as seen in Figure 8. More importantly, this easily extends to the dynamic case by using an optimal linear quadratic (LQ) state feedback controller [70] for linearizations of (9) at 0 and 100 watts, with relatively higher weights on BP and HR for the latter. Figure 9 compares HR and watts data versus (nonlinear) simulations of such controllers for two experiments. Thus the HR variability in Figure 7 (and BP in Figure 9) decreases with increasing watts because of straightforward and changing tradeoffs between metabolic overhead, ΔO2, HR, and BP, while their means increase.

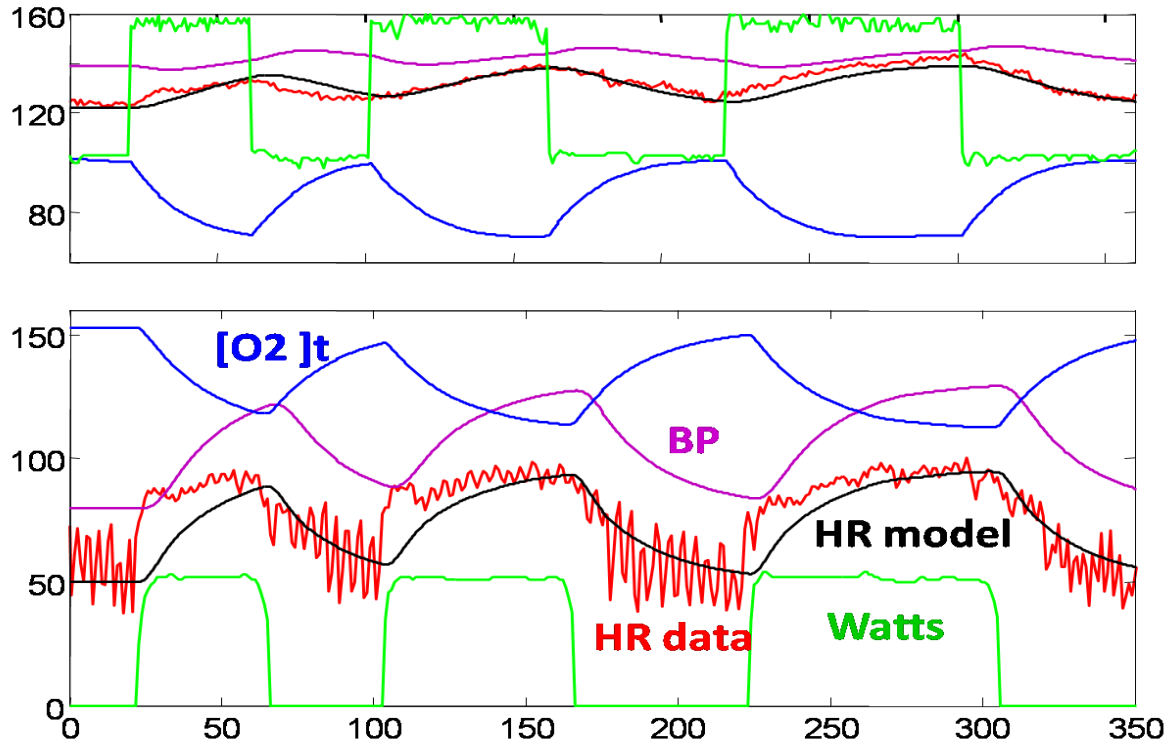


Figure 9 Optimal control model of heart rate (red) response of one subject to two different watts (green) demands, approximately square waves of 0-50 w (lower), 100-150w (upper)). For each data set, a first principle model is simulated with watts as input (green) and HR (black), blood pressure (purple) and tissue oxygen saturation (blue) as output. Breathing is natural (not shown).

Figure 10 sheds light on the nature of the high frequency HRV with two experiments at constant watts of 0 and 50 and a frequency sweep in breathing. For each level, HR variability is captured with a simple 2 state, 5 parameter linear model $\Delta h(t) = a_1 h(t) + b_1 v(t) + x(t)$, $\Delta x(t) = a_2 x(t) + b_2 v(t) + c$ where v is measured ventilation flow rate, x is an internal state, and parameters depend on watts. Despite lower breath magnitude the HR response is larger and faster at 0w than 50w. This is consistent with the trends in the other Figures, and simple global nonlinear "black box" models with both watts and breath inputs can equally fit all the data. Thus in these controlled experiments, even large HRV and changes in HRV can be explained as causal dynamic response to changing loads and pulsatile breathing, not chaos, but that the dependence of HRV on watts level is intrinsically nonlinear. Of course, our goal is not model fitting but physiological mechanisms, and we have extensively explored approaches to this. The above "black box" models are invaluable however since they clarify what signals are necessary to complete a model. Figure 7 shows that large slow fluctuations need explicit models of watts forcing, as do Figure 10 and pulsatile breathing. There is extensive literature on the former, making it an obvious starting point, and much less on the latter, which will be addressed next in our research.

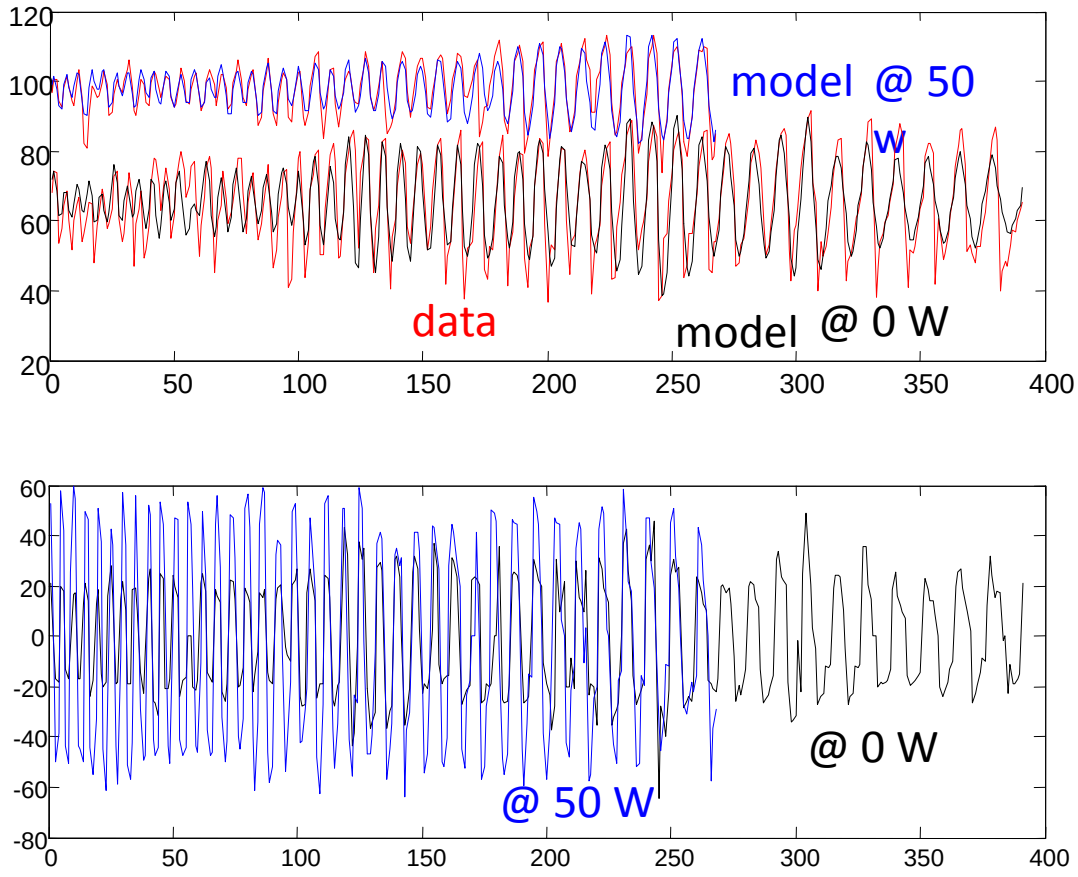


Figure 10 Frequency sweep in ventilation flow rate (lower plot) with fixed demand of 0w (black) and 50w (blue). Subject varied breathing to follow a preprogrammed frequency sweep spanning the natural breath frequencies at these watt levels. Breath flow magnitude was larger at 50w and the subject was unable to breathe slowly enough to complete the entire frequency sweep. For each data set, a second order linear model was fit with flow rate input (lower plot) and HR output (upper, data in red). Simulations are in upper plot for 0w (black) and 50w (blue).

Our initial approach has been to model the "plumbing and chemistry" of gases and pressure in physiological compartments that are involved in delivering power while maintaining brain homeostasis for gases and pressures. Then like for glycolysis, the tradeoffs that *any* controller must face provides a simple, rigorous, and mechanistic explanation for longstanding mysteries in cellular and physiological dynamics, while opening up whole new research directions. The more complex time series need black box models to clarify their causal relationships, and we have already found similar models for gases and BP signals, but both measurement and mechanism are more complex. The HRV tradeoffs already involve more complex mechanisms in (9) than in (1), but both models need expansion to include control of redox (and CO2), and the mechanistic effects of "internal" noise due to stochastics at the molecular level or pulsatile breathing and beating. The HRV tradeoffs use *less* sophisticated theory in Figure 8 and Figure 9 than in Figure 1, so there are additional tradeoffs from (5) in HR and particular BP control. These would be depend on the mechanisms implementing neuro-endocrine control, which are least well understood, and this is a major focus of our future research. Fortunately, the blend of black box, mechanististic, and optimal control models used here seems perfectly suited to this task as demonstrated for HRV. Such "grey box" modeling shows that high HRV at low watts results from an inevitable tradeoff between controlling gases and pressures to external watts demands versus response to "internal" noise such as pulsatile breathing.

Clinical implications

We are pursuing causal and mechanistic modeling of healthy cardiopulmonary control with the ultimate aim of developing alert algorithms for patient monitoring and diagnosis based on control theory. The indicator of an onset can potentially link to a specific physiological cause of failure, which leads to clinical context for clinicians to intervene precisely and early before serious damage or death. Monitoring has been a constant feature for anesthetized patients in the operating room as well as for patients in critical care units [68]. It is also utilized when cardiologists employ periods of continuous remote monitoring to evaluate patients for a variety of clinical conditions. Clinicians are beginning to realize the value of continuous monitoring in a variety of new settings such as medical-surgical hospital units and for home health care. As monitoring does become more ubiquitous, as well as more essential to a variety of care processes, it makes sense to fully leverage, analyze and maximize the information provided rather than to simply provide raw data for the clinical eye ball analysis that (even early twentieth century medicine) could provide.

The essence of traditional monitoring lies in the detection of simple anomalies such as a low blood pressure or a high heart rate. These vital sign values are the proxies for the characterization of the global physiological state and have progressed very little in analytic state over the past century. What we propose is an attempt to determine the biological factors that control the magnitude, directionality and dynamics of these values in a manner that will fundamentally alter the way clinicians view and use them. This will involve the ultimate creation of monitoring systems that *monitor the control elements* as well as the *controlled* elements (the vital sign values). This will provide clinicians with a wider window for timely interventions when dysfunction is still remediable and perhaps even more importantly, may provide fundamental new approaches to clinical diagnosis and therapeutics. There is an enormous opportunity here to create a new clinical paradigm that leverages technology to optimize elements of information that we already possess but whose utility is not even close to being maximized.

Selected Publications

[1] D. Alderson, J. Doyle. Contrasting views of complexity and their implications for network-centric infrastructures. *IEEE Trans Syst Man Cybern A: Systems And Humans* 40(4):839-852 (2010).

[2] M. Chiang, SH Low, AR Calderbank, JC Doyle, Layering As Optimization Decomposition, *PROCEEDINGS OF THE IEEE*, Volume: 95 Issue: 1 Jan 2007

[3] J. Lavaei, J. C. Doyle and S. Low. Congestion Control Algorithms from Optimal Control Perspective, *Proc IEEE Conference on Decision and Control*, 2009.

[4] J. Lavaei, JC Doyle, SH Low, Utility Functionals Associated With Available Congestion Control Algorithms, *IEEE International Conference on Computer Communications (INFOCOM)*, San Diego, California, 2010

[5] J. Lavaei**,** A Babakhani, A Hajimiri, J Doyle, Solving Large-Scale Hybrid Circuit-Antenna Problems, To appear in *IEEE Transactions on Circuits and Systems I*, 2010

[6] A. Babakhani, J. Lavaei, J. C. Doyle and A. Hajimiri, Finding Globally Optimum Solutions in Antenna Optimization Problems, to appear, *IEEE International Symposium on Antennas and Propagation*, 2010.

[7] J. Lavaei, A. Babakhani, A. Hajimiri and J. C. Doyle. Programmable Antenna Design Using Convex Optimization, *Proc 19th International Symposium on Mathematical Theory of Networks and Systems*, 2010.

[8] J. Lavaei, A. Babakhani, A. Hajimiri and J. C. Doyle. A Study of Near-Field Direct Antenna Modulation Systems Using Convex Optimization, *Proc American Control Conference*, 2010.

[9] J. Lavaei, A. Babakhani, A. Hajimiri and J. C. Doyle. Solving Large-Scale Linear Circuit Problems via Convex Optimization, *Proc IEEE Conference on Decision and Control*, 2009.

[10] J. Lavaei, A. Babakhani, A. Hajimiri and J. C. Doyle "'Passively Controllable Smart Antennas," to appear in *IEEE Global Communications Conference (GLOBECOM)*, Miami, Florida, 2010

[11] L. Chen, T. Cui and S. H. Low. A Game-theoretic Framework for Medium Access Control. *IEEE Journal of Selected Areas in Communications*, 2008.

[12] L. Chen, S. H. Low and J. C. Doyle. Random Access Game and Medium Access Control Design, *IEEE/ACM Transactions on Networking*, 2010.

[13] L. Chen, S. H. Low and J. C. Doyle. Cross-Layer Design in Multihop Wireless Networks (Invited), to appear, Special Issue on *Wireless for the Future Internet*, *Computer Networks Journal*, 2010.

[14] L. Chen, T. Ho, M. Chiang, S. H. Low, and J. C. Doyle. Congestion Control for Multicast Flows with Network Coding, under review for *IEEE Transactions on Information Theory*.

[15] T. Cui, L. Chen and T. Ho. Energy Efficient Opportunistic Network Coding for Wireless Networks. Proc. IEEE INFOCOM, Phoenix, AZ, Apr 2008.

[16] T. Cui, L. Chen and T. Ho, On Distributed Scheduling in Wireless Networks Exploiting Broadcast and Network Coding, *IEEE Transactions on Communications*, 2010.

[17] J. Doyle, M. Csete. Rules of engagement. *Nature* 446(7138):860 (2007).

[18] H. Kurata, H. El-Samad, R. Iwasaki, H. Ohtake, J. Doyle, et al. Module-based analysis of robustness tradeoffs in the heat shock response system. *PLoS Comput Biol* 2(7):e59 (2006).

[19] M. Csete, J. Doyle. Reverse engineering of biological complexity. *Science* 295(5560):1664-1669 (2002).

[20] JC Doyle, M. E. Csete, and L. Caporale, An engineering perspective: The implicit protocols, in *The Implicit Genome*, L. Caporale, Ed. London, U.K.: Oxford Univ. Press, 2007

[21] JC. Doyle, D. Alderson, L. Li, S. Low, M. Roughan, S. Shalunov,R. Tanaka, and W. Willinger, The 'robust yet fragile' nature of the Internet, Proc. Nat. Acad. Sci. U.S.A., vol. 102, no. 41, pp. 14 497–14 502,Oct. 2005.

[22] W. Willinger, D. L. Alderson, and J. C. Doyle, Mathematics and the Internet: A source of enormous confusion and great potential, *Notices AMS*, vol. 56, no. 5, pp. 586–599, May 2009.

[23] F. Chandra, G. Buzi, J. Doyle. Linear Control Analysis of the Autocatalytic Glycolysis System. *Proc American Control Conf:* 319-324 (2009).

[24] F. Chandra, G. Buzi, J. Doyle. Glycolytic Oscillations and Limits on Robust Efficiency, *Science*, Vol 333, pp 187-192

[25] H. Sandberg, J. Delvenne, J. Doyle. On lossless approximations, the Fluctuation-Dissipation Theorem, and limitations of measurements. , *IEEE Trans Auto Control*, Feb 2011.

[26] N. Martins, M. Dahleh, J. Doyle. Fundamental Limitations of Disturbance Attenuation in the Presence of Side Information. *IEEE Trans Auto Control* 52(1):56-66 (2007).

[27] D. Gayme, B. McKeon, A. Papachristodoulou, and J.C. Doyle. 2D/3C model of large scale structures in turbulence in plane Couette flow, *Proceedings of the 6th International Symposium on Turbulence and Shear Flow Phenomenon*, Seoul, Korea, June 2009.

[28] Gayme DF, McKeon BJ, Papachristodoulou P, Bamieh B, Doyle JC (2010) A streamwise constant model of turbulence in plane Couette flow, *J Fluid Mech*, vol 665, pp 99-119

[29] D. Gayme, B. McKeon, B. Bamieh, A. Papachristodoulou and J. C. Doyle. Amplification and nonlinear mechanisms in plane Couette Flow, *Physics of Fluids*, in press (published online 17 June 2011).
[30] D. Gayme, J. C. Doyle, A. Papachristodoulou, B. Bamieh and B. McKeon. Robustness and subcritical transition in wall-bounded shear flows, *Under preparation.*
[31] D. Gayme, *A robust control approach to understanding nonlinear mechanisms in shear flow turbulence*, 2010, PhD thesis, Control and Dynamical Systems, Caltech, Pasadena, USA.
[32] G. Buzi, U. Topcu, J. Doyle. Quantitative nonlinear analysis of autocatalytic pathways with applications to glycolysis. *Proc American Control Conf*: 3592–3597 (2010).
[33] G. Buzi, U. Topcu, J. Doyle. Compositional Analysis of the Autocatalytic Networks in Biology. *Proc American Control Conf* (2010).
[34] G. Buzi and J. Doyle. Topological Tradeoffs in Autocatalytic Metabolic Pathways. To appear in *Proc 49th IEEE Conf Decision and Control* (2010).
[35] J. Doyle, B. Francis, A. Tannenbaum. *Feedback Control Theory*. Macmillan Press, New York, NY (1992).
[36] F. A. Chandra, B. Genti and J. C. Doyle. Theoretical Performance Limit of the Autocatalytic Glycolysis System, *Proceedings of the 3rd Q-bio Conference on Cellular Information*, 2009.
[37] G. Buzi, U. Topcu and J. C. Doyle. Analysis of autocatalytic networks in biology, submitted, *Automatica*, 2010.
[38] N. Motee, F. Chandra, B. Bamieh, M. Khammash and J.C. Doyle. Performance limitations in autocatalytic networks in biology, To appear in *Proc IEEE Conference on Decision and Control* (2010).
[39] N. Motee and A. Jadbabaie, Distributed Multi-Parametric Quadratic Programming, IEEE Transactions on Automatic Control, Vol. 54 (10): 2279-2289, October 2009.
[40] N. Motee and A. Jadbabaie. Approximation methods and spatial interpolation in distributed control systems, Proc American Control Conference, St. Louis, Pages: 860-866, June 2009.
[41] N. Motee, B. Bamieh and M. Khammash, Stability analysis of a class of biological network models. Proc American Control Conference, Pages: 5936-5941, June 2010.
[42] N. Motee, B. Bamieh and M. Khammash. Model reduction of polynomial dynamical systems using differential algebra. To appear in *Proc IEEE Conference on Decision and Control* (2010).
[43] N. Motee, A. Jadbabaie and G. Pappas. Path planning for multiple robots: an alternative duality approach, Proc American Control Conference, Baltimore, Pages: 1611-1616, June 2010.
[44] N. Motee, A. Jadbabaie and G. Pappas. A duality approach to path planning for multiple robots, Proc of International Conference on Robotics and Automation, Pages: 935-940, May 2010.
[45] N. Motee, A. Jadbabaie, and G. Pappas. A duality theory for path planning for multiple autonomous vehicles, submitted, IEEE Transactions on Robotics, 2010.
[46] N. Motee, A. Ahmadzadeh, A. Jadbabaie and G. Pappas. Elastic Multi-Particle Systems for Multi-Vehicle Path Planning, submitted, IEEE Transactions on Automatic Control, 2010.
[47] J. Lavaei, S. Sojoudi and R. M. Murray. Simple Delay-Based Implementation of Continuous-Time Controllers, *Proc American Control Conference*, 2010.
[48] J. Lavaei, S. Sojoudi and R. M. Murray. Synthesis of Embedded Control Systems with High Sampling Frequencies, submitted, *IEEE Conference on Decision and Control*, 2010.
[49] J. Lavaei, S. Sojoudi and R. M. Murray. Delay-Based Controller Design for Continuous-Time and Hybrid Applications, under submission to *IEEE Transactions on Automatic Control*.
[50] J. Lavaei and R. M. Murray. On Quantized Consensus by Means of Gossip Algorithm--Part I: Convergence Proof, *Proc American Control Conference*, 2009.
[51] J. Lavaei and R. M. Murray. On Quantized Consensus by Means of Gossip Algorithm--Part II: Convergence Time. *Proc American Control Conference*, 2009.
[52] J. Lavaei and R. M. Murray. Quantized Consensus via Adaptive Stochastic Gossip Algorithm, submitted, *Proc IEEE Conference on Decision and Control*, 2009.
[53] J. Doyle, B. Recht, N. Li, J. Cruz, S. Sojoudi and S. Shien. Control Theoretic Perspective in Understanding Physiological Tradeoffs, in preparation, 2010.
[54] A. Lamperski. A neuro-inspired method for data rate limited control, submitted, *IEEE Conference on Decision and Control*, 2010.
[55] N. Li, J. R. Marden, and J. Shamma, Learning Approaches to the Witsenhausen Counterexample from a View of Potential Games, *Proceedings of IEEE Conference on Decision and Control*, 2009.
[56] N. Li and J. R. Marden, Designing Games to handle Coupled Constraints, *Proceedings of IEEE Conference on Decision and Control*, 2010.
[57] A. Lamperski, Representations and algorithms for finite-state bisimulations of linear discrete-time control systems, *Proceedings of IEEE Conference on Decision and Control*, 2009.

[58] A. Lamperski and Aaron D. Ames. Local zeno stability theory, submitted, *IEEE Transactions on Automatic Control*, 2010.
[59] A. Lamperski and A. D. Ames. On the existence of Zeno behavior in hybrid systems with non-isolated Zeno equilibra. *Proceedings of IEEE Conference on Decision and Control*, 2008.
[60] A. Lamperski and A. D. Ames. Sufficient conditions for Zeno behavior in Lagrangian hybrid systems. *In Hybrid Systems: Computation and Control, LNCS*, 2008.
[61] J. Lee, S. N. Sponberg, O. Y. Loh, A. Lamperski, R. J. Full, and N. J. Cowan. Templates and anchors for antenna-based wall following in cockroaches and robots. *IEEE Transactions on Robotics (Special Issue on Bio-Robotics)*, 5(1):130–143, 2008.
[62] T. Wongpiromsarn, S. Mitra, R. M. Murray, and A. Lamperski. Periodically controlled hybrid systems: Verifying a controller for an autonomous vehicle. In *Hybrid Systems: Computation and Control*, LNCS, 2009.
[63] T. Wongpiromsarn, S. Mitra, R. M. Murray, and A. Lamperski. Verification of periodically controlled hybrid systems: Application to an autonomous vehicle, to appear, *ACM Transactions on Embedded Computing Systems*, 2010.
[64] L. Chen, N. Li, and S. H. Low. On the interaction between load balancing and speed scaling. Under review for The 31st International Conference on Distributed Computing Systems (ICDCS 2011).
[65] JC Doyle**,** ME Csete (2011) Architecture, Constraints, and Behavior, *P Natl Acad Sci USA*, in press, available online

Other references

[66] J. Day, *Patterns in Network Architecture: A Return to Fundamentals*, Prentice Hall, 2008
[67] C.A. Mead, *Collective Electrodynamics: Quantum Foundations of Electromagnetism*, MIT Press, 2000
[68] David Stone, MD, private communication
[69] M. J. Cohen, F. A. Chandra, M. T. Ganter, J. Pittet, J. C. Doyle, A. Arkin A and K. Brohi, The Coagulation Network: Topology, Biology, and Evolution, submitted, *Critical Care Medicine*.
[70] K. Astrom, R. Murray. Feedback Systems: An Introduction for Scientists and Engineers. Princeton University Press, Princeton, NJ (2008).
[71] P. Richard. The rhythm of yeast. *FEMS Microb Rev*, vol. 27, pp. 547-557 (2003).
[72] B. Teusink, B. Bakker, H. Westerhoff. Control of frequency and amplitudes is shared by all enzymes in three models for yeast glycolytic oscillations. *Biochim. Biophys. Acta* 1275(3):204-12 (1996).
[73] M. Bier, B. Teusink, B. Kholodenko, H. Westerhoff. Control analysis of glycolytic oscillations. *Biophys Chem* 62(1-3):15-24 (1996).
[74] F. Hynne, S. Danø, P. Sørensen. Full-scale model of glycolysis in Saccharomyces cerevisiae. *Biophys Chem* 94(1-2):121-63 (2001).
[75] A. Goldbeter. *Biochemical Oscillations and Cellular Rhythms*. Cambridge University Press, Cambridge (1996).
[76] E. Sel'kov. Stabilization of energy charge, generation of oscillations and multiple steady states in energy metabolism as a result of purely stoichiometric regulation. *Eur J Biochem* 59:151-157 (1975).
[77] A. Betz, B. Chance. Phase relationship of glycolytic intermediates in yeast cells with oscillatory metabolic control. *Arch Biochem Biophys* 109(3):585-594 (1965).
[78] M. Savageau, P. Coelho, R. Fasani, et al. Phenotypes and tolerances in the design space of biochemical systems. *PNAS* 106(16):6435-6440 (2009).
[79] Y. Termonia, J. Ross. Oscillations and control features in glycolysis: Numerical analysis of a comprehensive model. *PNAS* 78(5):2952-2956 (1981).
[80] B. Chernyak, O. Pletjushkina, D. Izyumov, et al. Bioenergetics and death. *Biochemistry (Moscow)* 70(2):240-245 (2005).
[81] A. Poulsen, M. Petersen, L. Olsen. Single cell studies and simulation of cell–cell interactions using oscillating glycolysis in yeast cells. *Biophys Chem* 125(2-3):275-280 (2007).
[82] A. Ståhlberg, K. Elbing, J. Andrade-Garda, et al. Multiway real-time PCR gene expression profiling in yeast *Saccarhomyces cerevisiae* reveals altered transcriptional response of ADH-genes to glucose stimuli. *BMC Genomics* 9:170-185 (2008).
[83] I. Lestas, G. Vinnicombe, J. Paulsson. Fundamental limits on the suppression of molecular fluctuations. *Nature* 467(7312):174-178 (2010).
[84] P. Iglesias, B. Ingalls. *Control Theory and Systems Biology*. The MIT Press, Cambridge, MA (2010).
[85] E. Dessaud, V. Ribes, N. Balaskas, et al. Dynamic assignment and maintenance of positional identity in the ventral neural tube by the morphogen Sonic hedgehog. *PLoS Biol* 8(6):e1000382 (2010).

[86] L. Glass. Synchronization and rhythmic processes in physiology. *Nature* 410:277-284 (2001).
[87] L. Glass. Chaos and Heart Rate Variability. *J. Cardiovasc. Electrophysiol.* 10: 1358-1360 (1999).
[88] A. J. Camm, et al. Heart rate variability: standards of measurements, physiological interpretation, and clinical use. *Circulation* 93: 1043-1065 (1996).
[89] C. Poon, N. Holstein-Rathlou, E. Agner. Lack of evidence for low-dimensional chaos in heart rate variability, *J. Cardiovasc. Electrophysiol.* 5: 591-601 (1994).
[90] M. Costa, et al. No evidence of chaos in the heart rate variability of normal and cardiac transplant human subjects. *J. Cardiovasc. Electrophysiol.* 10: 1350-1357 (1999).
[91] P. Ivanov, et al. Multifractality in human heartbeat dynamics. *Nature* 399: 461-465 (1999).
[92] M. Costa, A. Goldberger, C. Peng. Multiscale Entropy Analysis of Complex Physiologic Time Series. *Physical Review Letters.* 89(6):068102(4) (2002).
[93] D. Chialvo, R. Gilmour Jr, J. Jalife, Low dimensional chaos in cardiac tissue. *Nature* 345:653-657 (1990).
[94] P. Ivanov, et al. Scaling behavior of heartbeat intervals obtained by wavelet-based time-series analysis. *Nature* 383: 323-327 (1996).
[95] P. Ivanov, et al. Multifractality in human heartbeat dynamics. *Nature* 399: 461-465 (1999).
[96] K. Wasserman, J. Hansen, D. Sue, W. Stringer, B. Whipp, *Principle of exercise testing and interpretation (Fourth Edition)*. Lippincott Williams & Wilkins, Philadelphia, PA, US (2005).
[97] G. Brooks, T. Fahey, K. Baldwin, *Exercise Physiology: Human Bioengergetics and Its Applications (Fourth Edition)*. Mc Graw Hill,US (2004)
[98] F. Hoppensteadt, C. Peskin. *Modeling and Simulation in Medicine and the Life Sciences (Second Edition)*. Texts in applied Mathematics 10, Springer, US (2002)
[99] J. Batzel, F. Kappel, D. Schneditz, H. Tran, *Cardiovascular and Respiratory Systems: Modeling, analysis, and Control*. SIAM series on Frontiers in Applied Mathematics, (2006).
[100] F. Grodins, J. Buell, A. Bart, *Mathematical Analysis and Digital Simulation of the Respiratory Control System*, Rand Corp, Santa Mocia, CA, US (1967)
[101] L. Ljung. *System Identification: Theory for the User (Second Edition*), PTR Prentice Hall, Upper Saddle River, NJ, US (1999).
[102] Stringer, J. Hansen, K. Wasseman. Cardiac output estimated non-invasively from oxygen uptake during exercise. *J Appl Physiol* 82: 908-912 (1997).
[103] L. Rowell. *Human Cardiovascular Control*, Oxford University Press, (1993).
[104] J. Hirsch, B. Bishop, Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate. *Am. J. Physiol.* 241 *(Heart Circ. Physiol.* 10*):* 620-629 (1981).
[105] Guyton, J. Hall. *Guyton and Hall Textbook of Medical Physiology (11th Edition),* Elsevier Saunders, Philadelphia, PA, US (2011)
[106] J. Taylor, C. Myers, J. Halliwill, H. Seidel, and D. Eckberg. Sympathetic restraint of respiratory sinus arrhythmia: implications for vagal-cardiac tone assessment in humans. *Am J Physiol Heart Circ Physiol* 280(6): H2804 - H2814 (2001).
[107] K. Brack, J. Coote and A. Ng, Interaction between direct sympathetic and vagus nerve stimulation on heart rate in the isolated rabbit heart. *Exp Physiol* 89: 128–139 (2004).
[108] H. Chiu, T. Kao. A mathematical model for autonomic control of heart rate variation. *IEEE Eng Med Biol* 20: 69–76 (2001).
[109] J. Elghozi, C. Julien. Sympathetic control of short-term heart rate variability and its pharmacological modulation. *Fundam Clin Pharmacol* 21:337–347 (2007).